

David C. Schultz, Ph. D.

AD _____

Award Number: DAMD17-98-1-8269

TITLE: Analysis BAP-1 as a Ubiquitin Hydrolase in the BRCA1 Pathway

PRINCIPAL INVESTIGATOR: David C. Schultz, Ph.D.

CONTRACTING ORGANIZATION: The Wistar Institute
Philadelphia, Pennsylvania 19104

REPORT DATE: October 2000

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
Distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20010427 089

Form Approved
OMB No. 074-0188

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE October 2000	3. REPORT TYPE AND DATES COVERED Annual Summary (1 Oct 99 - 30 Sep 00)
4. TITLE AND SUBTITLE Analysis BAP-1 as a Ubiquitin Hydrolase in the BRCA1 Pathway		5. FUNDING NUMBERS DAMD17-98-1-8269
6. AUTHOR(S) David C. Schultz, Ph.D.		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The Wistar Institute Philadelphia, Pennsylvania 19104		8. PERFORMING ORGANIZATION REPORT NUMBER
E-MAIL: dshultz@wistar.upenn.edu		
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012		10. SPONSORING / MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES This report contains colored photos		
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; Distribution unlimited		12b. DISTRIBUTION CODE
13. ABSTRACT (Maximum 200 Words) The recurrent theme that has emerged from the structural study of proteins is that multiple independently folded globular domains often cooperate in macromolecular recognition. These domains are often recognizable by conserved signature amino acid sequence motifs and their spatial organization within a novel protein is often the first clue to its biochemical function. The PHD finger and bromodomain are two highly conserved protein motifs found in proteins with transcriptional regulatory functions. Optimal transcriptional repression by KAP-1, a universal corepressor for the KRAB-ZFP superfamily of transcriptional repressors, requires an intact PHD finger and bromodomain at its COOH-terminus. Naturally occurring mutations in the PHD finger contribute to a variety of human diseases including cancer. Site-directed mutations in the KAP-1 PHD finger or the bromodomain specifically disrupt the repression function of this bi-partite domain. We have determined a solution structure for the PHD finger of KAP-1. Our studies reveal that the PHD domain binds zinc in a cross-brace topology reminiscent of RING domains. Mechanistically, the KAP-1 PHD and bromodomains are required to mediate an interaction with the Mi-2 α subunit of the NuRD <i>in vivo</i> as determined by co-immunoprecipitation. These data indicate that the PHD finger and bromodomain provide an interface for protein-protein interactions. Moreover, the data suggest the KRAB-ZFP:KAP-1 repression complex functions to target the histone deacetylase activity of the NuRD complex to specific gene promoters <i>in vivo</i> . Recent reports have identified a novel KRAB-ZFP that associates with BRCA1 implicating KRAB:KAP-1 repression in hereditary breast cancer.		
14. SUBJECT TERMS Breast Cancer, Transcriptional repression, KRAB-ZFP, KAP-1, PHD finger		15. NUMBER OF PAGES 118
		16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified
		20. LIMITATION OF ABSTRACT Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

Cover.....	1
SF 298.....	2
Introduction.....	4
Body.....	4
Key Research Accomplishments.....	10
Reportable Outcomes.....	10
Conclusions.....	11
References.....	11
Appendices.....	attached

Introduction:

The molecular basis of human disease is rapidly being defined. Genetic analyses of large affected kindreds have greatly facilitated the identification of a gene or set of genes when mutated is instrumental to the etiology of that particular disease. In some instances the clinical phenotype (i.e. UV sensitivity) presented by the patient may provide ideas as to which cellular pathway(s) (i.e. DNA repair) are affected. Contrary to this scenario, many patients lack a definable clinical phenotype, which can be associated with a cellular pathway, which may be affected. Nonetheless, once a new disease gene is identified, definition of its protein's function is the immediate interest in trying to understand why and how naturally occurring mutations in the gene perturb the function of its protein and confer disease susceptibility.

As we begin to enter the post-genome era, structure/function studies of newly cloned gene products will be described by many different paradigms. The recurrent theme that has emerged from the biochemical/structural study of proteins involved in signal transduction and transcriptional regulation is that multiple independently folded globular domains in these proteins often cooperate in macromolecular recognition. The presence of highly conserved amino acid signature motifs and their spatial arrangement within a novel protein can provide the first clue to the protein's function. Furthermore, the identification of interacting proteins has been beneficial in defining the biochemical function of a particular protein and/or the cellular pathway (i.e. cell signalling, transcription, protein degradation) in which it functions. Multi-disciplinary studies aimed at analyzing a protein's structure, biochemical function, and cellular pathway have been fruitful in understanding why naturally occurring mutations in a given protein confer susceptibility to disease. These studies have assisted in the design and development of better detection systems, disease diagnosis, and innovative therapeutic strategies for the intervention of disease. The **purpose** of this post-doctoral fellowship was to broaden my training in molecular biology/biochemistry, with an emphasis on understanding the biochemical functions of a protein and its physiologic role in the cell.

Current Research:

The fundamental characteristic of cells to detect extracellular stimuli and execute the appropriate response is critical for Metazoan development and homeostasis. Upon receiving external signals, finely orchestrated cascades transduce signals from the cell surface to the nucleus that culminate in the activation or repression of defined sets of genes which leads to integrated cellular responses such as proliferation, differentiation, and apoptosis. While core promoter elements facilitate the assembly of and orient the basal transcription machinery at the transcription start site, unique cis-regulatory elements upstream from the core promoter are recognized by sequence-specific transcription factors. These factors are frequently the penultimate targets of extracellular signals and influence the rate at which a given gene is transcribed. My **current research interests** have focused on structural and biochemical mechanisms of transcriptional repression conferred by superfamilies of sequence specific transcription factors.

Regulation of RNA polymerase II involves a complex interplay between DNA-protein interactions and protein-protein interactions. While the general transcription factors regulate the accurate initiation of transcription, the proteins that bind gene-specific DNA elements that are instrumental in either negatively or positively regulating the rate of transcription. The dominant theme that has emerged from the study of eukaryotic transcriptional regulatory proteins is that they are highly modular in architecture with independent, functionally-separable domains mediating nuclear localization, sequence-specific DNA binding, hetero- or homo-oligomerization, activation, and repression of transcription. Recently, much effort has been expended to understand how activation and repression domains transmit the signal for modulation of transcription from a DNA-bound protein

to the RNA synthesis machinery. Studies aimed at understanding the mechanisms of transcription repression have been greatly aided by the realization that the domains, which mediate repression, are often highly conserved amino-acid sequence motifs, linked to a common DNA binding domains. Examples of these domains include BTB/POZ, WRPW, SNAG, and KRAB. We have focused on the Krüppel-associated Box (KRAB) domain as a model system for the analysis of transcriptional repression in mammalian cells (1).

The KRAB (Krüppel associated box) domain is one example of an abundant amino acid sequence motif found at the NH₂-terminus of nearly one-third of all Krüppel/TFIIIA type C2H2 zinc finger protein (1). This highly conserved domain displays potent, DNA-binding dependent repression of transcription that requires the KAP-1 corepressor (2,3). The primary amino acid sequence of KAP-1 revealed the presence of several conserved signature motifs, including a RING finger, B boxes, a coiled-coil region (RBCC), which collectively form an integrated domain that is both necessary and sufficient to directly interact with the KRAB domain (4,5). The COOH-terminal sequence of KAP-1 revealed a conserved PHD finger and bromodomain (6,7). This particular spatial arrangement of motifs has defined an emerging family of transcriptional regulators, which include TIF1 α and TIF1 γ . All members of this protein family have been shown to repress transcription when tethered to DNA, and the mechanisms by which they repress transcription are currently being defined (8,9).

To identify regions of KAP-1 which are required for its transcriptional repression function we have engineered a series of KAP-1 deletion mutants with the heterologous GAL4 DNA binding domain to grossly map repression domains. In general, we have found that amino acids COOH-terminal to the RBCC region define at least two independent repression domains within KAP-1 that appear to be additive in nature in order to obtain maximal levels of transcriptional repression. One of these domains includes the recently defined HP1 binding domain of KAP-1 (10,11). Optimal repression of transcription was dependent upon the PHD finger and bromodomain. When autonomously tethered to DNA, these two motifs demonstrated a significant level of repression, suggesting that they can function as an independent repression domain. Neither the PHD finger nor the bromodomain could independently recapitulate this repression activity, indicating that these two motifs function cooperatively to repress transcription. Furthermore, expression of the segment encoding the PHD finger and bromodomain of KAP-1 relieved both GAL4-KRAB and GAL4-KAP-1 directed repression, suggesting that repression by this bipartite domain is mediated by at least one or more titratable effector molecules. The combination of these data suggests that the PHD finger and bromodomain of KAP-1 function as an integrated unit to facilitate repression of transcription by KAP-1. Moreover, these data implicate an additional repression mechanism that is independent of the recruitment of HP1 proteins.

Because the KAP-1 orthologues, TIF1 α and TIF1 γ , have been reported to function as transcriptional repressors, we compared the repression properties of the PHD finger and bromodomain of all three family members. To directly test this hypothesis, we constructed heterologous fusions between the GAL4 DNA binding domain and the PHD finger and bromodomains of TIF1 α , TIF1 γ , and WCRF180. In addition, we engineered chimeric proteins in which either the PHD finger or bromodomain of TIF1 α , TIF1 γ , hATRX, Mi-2 α , WCRF180, or hGCN5 were substituted for the KAP-1 sequences in order to address the specificity of the repression activity displayed by this bipartite domain. When tethered to DNA, the PHD finger and bromodomain of TIF1 α demonstrated a potential to activate transcription of the reporter template, while the PHD finger and bromodomain of TIF1 γ and WCRF180 each failed to demonstrate any appreciable potentiation of transcription. Heterologous substitution of the KAP-1 PHD finger or bromodomain with similar sequences from other

transcriptional regulators failed to completely complement the repression activity of the wild-type KAP-1 PHD finger and bromodomain. The combination of these data provides additional support that the PHD finger and bromodomain of KAP-1 function as an integrated functional unit to specifically repress transcription. The data also imply that not all PHD fingers and bromodomains are functionally equivalent, despite demonstrating homology in the primary amino acid sequence. Furthermore, the spatial conservation of these two motifs architecturally in proteins is not sufficient to infer a common function for this bipartite domain.

In order to correlate the molecular relationship between amino acids in the PHD finger and bromodomain and transcriptional repression, we have used a site-directed mutagenesis approach. All amino acid substitutions were made in the context of the heterologous GAL4-PHD finger/bromodomain fusion protein. The effect of each mutation on the repression properties of this bipartite domain was determined in DNA-template based assays. As an initial strategy, we have targeted amino acid residues in the PHD finger of KAP-1 which spatially parallel the position of naturally occurring mutations in hATRX that confer an inherited susceptibility to developmental defect (12). Second, we targeted amino acid residues that are strictly conserved among PHD fingers most closely related to KAP-1. The final scheme focused on the PHD finger involved mutating amino acids in the variable regions of the KAP-1 PHD finger to residues present in the PHD fingers of TIF1 α /TIF1 γ . In general, mutation of any of the putative zinc-chelating cysteine residues and several highly, conserved hydrophobic amino acids significantly disrupted the repression activity of this bipartite domain (Schultz et al, submitted, Capili et al, submitted).

To gain additional insights into the biochemical function of the PHD finger, we expressed the minimal KAP-1 PHD (amino acids 618-679) in *Escherichia coli* and purified the protein to homogeneity using affinity and size exclusion chromatographies. A combination of inductively coupled plasma (ICP) spectrometry, circular dichroism, and homonuclear/heteronuclear NMR techniques revealed that the KAP-1 PHD finger binds zinc in a cross-brace topology between anti-parallel β -strands reminiscent of RING domains (Capili et al submitted). To determine the effects that mutations impose on the structural properties of the PHD finger, we expressed the PHD protein containing mutations of three different types (hydrophobic core, zinc chelation, addition of an extra zinc ligand). Gel filtration, 1D NMR and CD studies on all three mutations indicated that, as expected, disruption of the hydrophobic core or of zinc binding resulted in unfolded proteins. ^1H NMR studies indicate that α -carbon resonances corresponding to secondary structural elements present in the wildtype proteins are absent in the mutants. In addition, there is poor dispersion of the resonances in the amide region of the spectra of all three mutants indicative of an unfolded protein. Further, there are few NOEs observed in 2D ^1H NOESY experiments again demonstrating that these proteins are largely unstructured. Thus, the ability of the KAP-1 PHD domain to bind zinc and maintain the hydrophobic core are both required for its proper folding and ability to repress transcription.

We also used the information provided by the solution structure of the pCAF bromodomain to design six independent mutations in each of the four α -helices and one mutation in the variable loop region between helix Z and helix A of the KAP-1 bromodomain (13). Mutations in both helix B and helix C significantly impaired transcriptional repression. When these mutations were mapped back onto the 3D structure, each is presumed to effect amino acids that contribute to the hydrophobic core and the stability of the structure. Interestingly, the mutation in the variable loop region between helix Z and helix A also displayed an intermediate effect on the repression potential of this domain, reaffirming the potential role for this region in macromolecular recognition. Overall, the data from this mutational analysis is consistent with the PHD finger and bromodomain of KAP-1 functioning

cooperatively, as independent mutations in either domain can significantly disrupt the intrinsic repression potential of this bipartite domain. One interpretation of this data is that these two motifs provide an interface for protein-protein interactions with downstream effectors of transcriptional repression.

To identify potential mechanisms of transcriptional repression directed by the KAP-1 PHD finger and bromodomain, a heterologous fusion between the LEXA DNA binding domain and the PHD finger and bromodomain of KAP-1 was used in a yeast two-hybrid screen. To confirm the specificity of any potential protein interactions identified in this screen, we utilized several of the amino acid substitutions engineered for the evaluation of the structure/function analysis of this bipartite domain. Four different gene products were recovered based on coactivation of both the integrated LEXA responsive *His3* and *LacZ* reporter genes. The nucleotide sequence of three, independent transformants of one gene product indicated that the sequence was identical to the 3' nucleotide sequence of the dermatomyositis-specific autoantigen, Mi-2 α /CHD3 (14). Upon reintroduction into yeast, this GAL4 activation domain fusion protein displayed robust ability to activate the *LacZ* reporter with the wild-type LEXA-KAP-1 fusion protein, but failed to activate the reporter with other irrelevant test baits. Two PHD finger mutants that demonstrated nearly wild-type levels of transcriptional repression activity also displayed the ability to activate the *LacZ* reporter. In contrast, zinc chelating or hydrophobic core mutations in the PHD finger, which possess impaired repression activity, either failed to activate these reporters or activated with reduced efficiency, respectively. This putative protein-protein interaction was also observed to be dependent on a functional bromodomain. These observations are completely consistent with the transcriptional effects observed in transient transfection assays in which the repression activity was dependent upon both domains. Furthermore, mutations in either the PHD finger or the bromodomain were independently sufficient to ablate the repression function of this bipartite domain. The combination of these data strengthens the argument that the PHD finger and bromodomain of KAP-1 function as an integrated functional unit, which provides an interface for protein-protein interactions with molecules that facilitate repression by the KRAB:KAP-1 complex.

The Mi-2 family of proteins has been described as an integral component of a high molecular weight multiprotein complex containing histone deacetylase activity (15,16). To verify an interaction *in vivo* between KAP-1 and Mi-2 α we used coimmunoprecipitation experiments of endogenous proteins from phosphocellulose fractionated HeLa nuclear extracts. Three independent antibodies against nonoverlapping antigens of KAP-1 specifically coprecipitated Mi-2 α , HDAC1, and RbAp48, three constitutive polypeptides of the NuRD complex. This result demonstrates an association between KAP-1 and Mi-2 α *in vivo*. To determine the specificity of this interaction *in vivo*, we transiently transfected 293 cells with Flag-tagged wild-type KAP-1 or a Flag-tagged delta PHD/bromodomain construct. Coimmunoprecipitation experiments revealed the specific precipitation of endogenous Mi-2 α with the wild-type KAP-1 protein, but not with the delta PHD/bromodomain KAP-1 protein. This data demonstrates that the association of KAP-1 with Mi-2 α specifically requires sequences in the PHD finger and bromodomain. In separate experiments, overexpression of the C-terminal amino acids of Mi-2 alpha can dominantly inhibit KRAB:KAP-1-mediated repression *in vivo*. Since Mi-2 has been biochemically purified as an integral component of a core histone deacetylase complex, these studies suggest that the KRAB:KAP-1 repression complex assists in targeting the biochemical activities of NuRD to specific promoters *in vivo*. Consistent with this hypothesis, addition of trichostatin A to transient transfection assays revealed that the transcriptional repression exerted by the PHD finger and bromodomain of KAP-1 is partially reversible, indicating the potential involvement of histone deacetylases in KAP-1 mediated repression (Schultz et al., submitted). These data provide evidence

that the KRAB-ZFP:KAP-1 repressor:corepressor complex functions to target histone deacetylases to KRAB-zfp target genes *in vivo*.

Future Directions:

The **goal** of this research fellowship was to further broaden my training in molecular biology/biochemistry with a particular focus on understanding the biochemical functions of novel gene products, which may have a role in the etiology of human disease. The findings of the current research have greatly fostered the development of new hypotheses to test in the future. As I prepare to establish an independent laboratory, the **focus** of my research will continue to use genetic and biochemical approaches to ask questions about the molecular relationship between the biochemical functions of a protein and its physiologic role in the cell. In particular, I am interested in **structural and biochemical mechanisms of transcriptional repression** conferred by superfamilies of mammalian sequence specific transcription factors. The **long-range goal** of this work is to understand how targeted recruitment of repressor:corepressor molecules facilitate transcriptional repression of chromatin templates. The **rationale** for this research is that once the mechanics of transcriptional repression of gene promoters at the chromatin level is known, new and innovative approaches, both genetic and pharmacologic, can be designed to reverse aberrant, epigenetic silencing of genes involved in a variety of diseases. This direction is particularly relevant to breast cancer, as several genes have been shown to be epigenetically silenced (i.e. E-cadherin) during mammary tumorigenesis. Furthermore, a recent report has described the association of a novel KRAB-ZFP with BRCA1 (17). This association implies a role for KAP-1 in BRCA1mediated transcriptional repression.

Our studies, which represent the first structural characterization of a PHD domain, will have widespread implications, as more than 400 independent proteins have been described to possess a PHD finger. Moreover, our structural analysis help explain the consequences of site-specific mutations in the PHD finger of the hATRX, AIRE1, and ING1 proteins which predispose individuals to developmental defects and cancer. The coordination of zinc in a cross-brace topology between anti-parallel β -strands reminiscent of RING domains invokes the possibility that the PHD finger may have RING-like biological activities. Recent studies have shown that some RING domains function as E3 ubiquitin ligases, components of the ubiquitin degradation pathway. Because of the similarity of the PHD finger to the RING, it is reasonable to hypothesize that the PHD finger may also function as an E3 for ubiquitin or any other ubiquitin-like molecule. We have initiated experiments to test whether the RING finger and PHD finger of KAP-1 or any other members of the TIF1 family of transcriptional regulators possess this E3 activity. Such studies may provide information as to how this unique family of proteins functions to regulate transcription, and moreover, may provide insights into how these proteins themselves are regulated. As an extension of our structural studies, we are attempting to solve the crystal structure of a PHD finger/bromodomain protein. Since our results indicate that the PHD finger and bromodomain of KAP-1 likely function as an integrated unit, these results are expected to provide new insights as to how this bipartite domain functions in transcriptional repression.

Second, I am interested in understanding the mechanics of transcriptional silencing of genes at single copy genomic loci. We are particularly well prepared to undertake the proposed research, because we have extensively defined the biochemical mechanisms of transcriptional repression utilized by the model KRAB:KAP-1 repressor:corepressor system, including heterochromatinization, histone deacetylases, and histone methylases. From these data, we postulate that transcriptional repression *in vivo* results from the ordered progression of independent molecular events. We will use

the well-defined KRAB:KAP-1 repression system to dissect the spatial and temporal order of these events at a heterologous-targeted transgene locus. Chromatin immunoprecipitation and micrococcal nuclease digestion experiments will be used to monitor the physical properties of the transcriptionally silent chromatin locus. Previously defined dominant mutations in the KAP-1 corepressor which affect protein interactions with effectors of repression will be useful in dissecting the role a given mechanism has in the efficiency of gene silencing by KAP-1. The results from these studies can be used as a framework to test gene silencing of endogenous KRAB-ZFP target genes. In particular, we will focus our efforts on ZBRK1, a BRCA1 associated KRAB-ZFP, and its role in the regulation of the GADD45 gene (17).

The proposed research for the final year of this fellowship is **innovative**, because its applies the well characterized KRAB:KAP-1 repression system to evaluate the function of the PHD finger, an abundant amino acid motif in putative transcriptional regulatory proteins, and specific effects of targeted transcriptional silencing at a single copy genomic loci. It is our **expectation** that the proposed research will define molecular events within promoter elements that occur during transcriptional silencing. These results will be **significant**, because the knowledge of transcriptional gene silencing mechanics will provide new targets for therapeutic intervention of *de novo* silenced genes in cancer. In addition, it is expected that the results of this proposal will fundamentally advance the field of covalent histone modifications and the biochemistry of transcriptional repression by the KRAB zinc finger protein superfamily.

Key research accomplishments

- Described an autonomous DNA-binding dependent repression activity conferred by the PHD finger and bromodomain of KAP-1.
- Confirmed an *in vivo* association between the endogenous corepressor KAP-1 and the Mi-2 α subunit of the histone deacetylase complex, NuRD.
- Demonstrated that the *in vivo* association between KAP-1 and Mi-2 α requires the PHD finger and bromodomain of KAP-1.
- Solved the first solution structure of the KAP-1 PHD finger.
- Using the KAP-1 PHD finger structure, characterized the chemical and physical effect of spatially conserved disease relevant mutations.

Reportable Outcomes:

Publications

Peng H, Begg GE, **Schultz DC**, Friedman JR, Jensen DE, Speicher DW, Rauscher III FJ. Reconstitution of the KRAB-KAP-1 Repressor Complex: A Model System for Defining the Molecular Anatomy of RING-B-box-Coiled-Coil Domain-Mediated Protein-Protein Interactions. *J Mol. Bio.* 305: 1139-1162 (2000).

Manuscripts

Schultz DC, Friedman JR, Rauscher III FJ Targeting Histone Deacetylase Complexes via KRAB-Zinc Finger Proteins: The PHD and Bromodomains of KAP-1 form a Cooperative Repression Domain that Recruits the Mi-2 α Subunit of the NuRD complex. (Submitted, *Genes and Development*, 2000)

Capili AD, **Schultz DC**, Rauscher III FJ, and Borden KLBSolution structure of the PHD domain from the KAP-1 corepressor: Structural determinants for PHD, RING, and LIM zinc-binding domains. (Submitted, *EMBO*, 2000)

Abstracts

***Schultz DC**, Jensen DE, Peng H, Rauscher FJ, III. The BRCA1 and KAP-1 RING fingers: model domains for studying the function of transcriptional coactivator and corepressor proteins. The Era of Hope Breast Cancer Meeting, June 9, 2000.

Schultz DC, Capili AD, Borden KL, Rauscher III FJ. The KRAB-ZFP:KAP-1 Repression System Recruits the Histone Deacetylase Complex, NuRD, via an Interaction Between the PHD and Bromodomains of KAP1 and the Mi-2 Subunit. (Submitted Kestone Symposia-C5, 2001)

- Indicates platform presentation given by David C. Schultz

Conclusions:

The support fellowship has provided the fundamental training in both molecular biology and biochemistry techniques. The training has allowed new hypotheses to be formulated, based on completed studies and current work, to extend this research in the next logical direction. Moreover the this fellowship training has provided the intellectual framework for the design of new research projects which address important questions pertaining to breast cancer.

References:

1. Bellefroid, E. J., Poncelet, D. A., Lecocq, P. J., Revelant, O., and Martial, J. A. (1991) *Proceedings of the National Academy of Sciences of the United States of America* **88**(9), 3608-12
2. Margolin, J. F., Friedman, J. R., Meyer, W. K., Vissing, H., Thiesen, H. J., and Rauscher, F. J. I. (1994) *Proceedings of the National Academy of Sciences of the United States of America* **91**(10), 4509-4513
3. Friedman, J. R., Fredericks, W. J., Jensen, D. E., Speicher, D. W., Huang, X. P., Neilson, E. G., and Rauscher, F. J. I. (1996) *Genes and Development* **10**(16), 2067-2078
4. Peng, H., Begg, G. E., Harper, S. L., Friedman, J. R., Speicher, D. W., and Rauscher, F. J. I. (2000) *J Biol Chem* **in press**
5. Peng, H., Begg, G. E., Schultz, D. C., Friedman, J. R., Jensen, D. E., Speicher, D. W., and Rauscher, F. J. I. 3rd. (2000) *J Mol Biol* **295**(5), 1139-1162
6. Aasland, R., Gibson, T. J., and Stewart, A. F. (1995) *Trends Biochem Sci* **20**(2), 56-59
7. Jeanmougin, F., Wurtz, J. M., Le Douarin, B., Chambon, P., and Losson, R. (1997) *Trends Biochem Sci* **22**(5), 151-153
8. Le Douarin, B., Nielsen, A. L., Garnier, J. M., Ichinose, H., Jeanmougin, F., Losson, R., and Chambon, P. (1996) *Embo J* **15**(23), 6701-6715
9. Venturini, L., You, J., Stadler, M., Galien, R., Lallemand, V., Koken, M. H., Mattei, M. G., Ganser, A., Chambon, P., Losson, R., and de The, H. (1999) *Oncogene* **18**(5), 1209-1217
10. Ryan, R. F., Schultz, D. C., Ayyanathan, K., Singh, P. B., Friedman, J. R., Fredericks, W. J., and Rauscher, F. J. I. 3rd. (1999) *Mol Cell Biol* **19**(6), 4366-4378
11. Lechner, M. S., Begg, G. E., Speicher, D. W., and Rauscher, F. J. I. (2000) *Mol Cell Biol* **in press**
12. Gibbons, R. J., Bachoo, S., Picketts, D. J., Aftimos, S., Asenbauer, B., Bergoffen, J., Berry, S. A., Dahl, N., Fryer, A., Keppler, K., Kurosawa, K., Levin, M. L., Masuno, M., Neri, G., Pierpont, M. E., Slaney, S. F., and Higgs, D. R. (1997) *Nat Genet* **17**(2), 146-148
13. Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M. (1999) *Nature* **399**(6735), 491-496
14. Woodage, T., Basrai, M. A., Baxevanis, A. D., Hieter, P., and Collins, F. S. (1997) *Proc Natl Acad Sci U S A* **94**(21), 11472-11477
15. Wade, P. A., Jones, P. L., Vermaak, D., and Wolffe, A. P. (1998) *Curr Biol* **8**(14), 843-846
16. Zhang, Y., LeRoy, G., Seelig, H. P., Lane, W. S., and Reinberg, D. (1998) *Cell* **95**(2), 279-289
17. Zheng, L., Hongyi, P., Shang, L., Andrea, F.-N., Phang-Lang, C., G, B. T., and Wen-Hwa, L. (2000) *Molecular Cell* **6**, 757-768

**THE BRCA1 AND KAP-1 RING FINGERS: MODEL DOMAINS FOR
STUDYING THE FUNCTIONS OF TRANSCRIPTIONAL COACTIVATOR AND
COREPRESSOR PROTEINS**

**David C. Schultz, David E. Jensen, Hongzhuang Peng, and
Frank J. Rauscher, III**

The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104

Email: dshultz@wistar.upenn.edu

The recurrent theme that has emerged from the structural study of proteins involved in signal transduction and transcriptional regulation is that multiple independently folded globular domains in these proteins often cooperate in macromolecular recognition. The presence of highly conserved signature amino acid sequence motifs and their spatial organization within a novel protein is often the first clue to its biochemical function. Furthermore, the identification of interacting proteins has been beneficial in defining the biochemical function of a particular protein and/or the cellular pathway (i.e. cell signaling, transcription, DNA repair, protein degradation) in which it functions. We have initially used the RING finger motif as a model domain for studying protein function. The RING finger motif is a very abundant sequence motif in the protein database. Upon cloning of *BRCA1*, the most recognizable signature motif in the amino acid sequence was a RING finger at the N-terminus. Moreover, genetic analysis of breast cancer kindred's identified naturally occurring germline mutations that lead to single amino acid substitutions of highly conserved residues in the RING finger and predispose individuals to disease. We have identified the novel protein, BAP-1, that specifically interacts with the RING finger of *BRCA1*, but not naturally occurring RING mutants. These initial studies have expanded to understanding the function of the RING finger motif in the context of a tripartite RBCC domain (RING-B-box-coiled-coil), a highly conserved spatial arrangement of signature motifs found in a number of oncoproteins, including PML1 and MID1. We have found that each sub-domain of this tripartite motif of KAP-1, a transcriptional corepressor protein, is essential for interaction with its ligand, and each of these structural elements appears to be vital for proper protein oligomerization. These studies will have widespread implications in future studies aimed at determining the ligands for RBCC containing proteins. The PHD finger motif is a second autonomous signature motif, which significantly resembles the spatial arrangement of the RING finger. Similar to the RING finger, naturally occurring mutations in the PHD finger of the hATRX protein confer a susceptibility to developmental defects. In this regard, synthetically derived mutations in the PHD finger of KAP-1 disrupt its ability to optimally repress transcription. Ultimately, the detailed analysis of a protein's structure, biochemical function, or cellular pathway may assist in the design and development of better therapeutic strategies for the intervention of disease.

The U. S. Army Medical Research and Material Command under DAMD 17-98-1-8269 supported this work.

**The KRAB-ZFP:KAP-1 Repression System
Recruits the Histone Deacetylase Complex, NuRD, via an
Interaction Between the PHD and Bromodomains of KAP1
and the Mi-2 Subunit.**

David C. Schultz¹, Allan D. Capili², Katherine L. B. Borden²,
and Frank J. Rauscher III¹. ¹The Wistar Institute, Philadelphia
PA. ²Mount Sinai School of Medicine, New York, NY.

The PHD finger and bromodomain are two highly conserved protein motifs found in proteins with transcriptional regulatory functions. Naturally occurring mutations in the PHD finger contribute to a variety of human diseases including ATRX syndrome, myeloid leukemias, and autoimmune dysfunction. Optimal transcriptional repression by KAP-1, a universal corepressor for the KRAB-ZFP superfamily of transcriptional repressors, requires an intact PHD finger and bromodomain at its COOH-terminus. Site-directed mutations in the KAP-1 PHD finger or the bromodomain specifically disrupt the repression function of this bi-partite domain. We have determined a solution structure for the PHD finger of KAP-1. Our studies reveal that the PHD domain binds zinc in a cross-brace topology reminiscent of RING domains. Mechanistically, the KAP-1 PHD and bromodomains are required to mediate an interaction with the Mi-2 α subunit of the NuRD *in vivo* as determined by co-immunoprecipitation. These data indicate that the PHD finger and bromodomain provide an interface for protein-protein interactions. Moreover, the data suggest the KRAB-ZFP:KAP-1 repression complex functions to target the histone deacetylase activity of the NuRD complex to specific gene promoters *in vivo*. This work was funded in part by DAMD 17-98-1-8269 to DCS.



Reconstitution of the KRAB-KAP-1 Repressor Complex: A Model System for Defining the Molecular Anatomy of RING-B Box-coiled-coil Domain-mediated Protein-protein Interactions

Hongzhuang Peng, Gillian E. Begg, David C. Schultz, Josh R. Friedman
David E. Jensen, David W. Speicher and Frank J. Rauscher III

Reconstitution of the KRAB-KAP-1 Repressor Complex: A Model System for Defining the Molecular Anatomy of RING-B Box-coiled-coil Domain-mediated Protein-protein Interactions

Hongzhuang Peng, Gillian E. Begg, David C. Schultz, Josh R. Friedman
David E. Jensen, David W. Speicher and Frank J. Rauscher III*

The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA

The KRAB domain is a 75 amino acid residue transcriptional repression module commonly found in eukaryotic zinc-finger proteins. KRAB-mediated gene silencing requires binding to the corepressor KAP-1. The KRAB:KAP-1 interaction requires the RING-B box-coiled coil (RBCC) domain of KAP-1, which is a widely distributed motif, hypothesized to be a protein-protein interface. Little is known about RBCC-mediated ligand binding and the role of the individual sub-domains in recognition and specificity. We have addressed these issues by reconstituting and characterizing the KRAB:KAP-1-RBCC interaction using purified components. Our results show that KRAB binding to KAP-1 is direct and specific, as the related RBCC domains from TIF1 α and MID1 do not bind the KRAB domain. A combination of gel filtration, analytical ultracentrifugation, chemical cross-linking, non-denaturing gel electrophoresis, and site-directed mutagenesis techniques has revealed that the KAP-1-RBCC must oligomerize likely as a homo-trimer in order to bind the KRAB domain. The RING finger, B2 box, and coiled-coil region are required for oligomerization of KAP-1-RBCC and KRAB binding, as mutations in these domains concomitantly abolished these functions. KRAB domain binding stabilized the homo-oligomeric state of the KAP-1-RBCC as detected by chemical cross-linking and velocity sedimentation studies. Mutant KAP-1-RBCC molecules hetero-oligomerize with the wild-type KAP-1, but these complexes were inactive for KRAB binding, suggesting a potential dominant negative activity. Substitution of the coiled-coil region with heterologous dimerization, trimerization, or tetramerization domains failed to recapitulate KRAB domain binding. Chimeric KAP-1-RBCC proteins containing either the RING, RING-B box, or coiled-coil regions from MID1 also failed to bind the KRAB domain. The KAP-1-RBCC mediates a highly specific, direct interaction with the KRAB domain, and it appears to function as an integrated, possibly cooperative structural unit wherein each sub-domain contributes to oligomerization and/or ligand recognition. These observations provide the first principles for RBCC domain-mediated protein-protein interaction and have implications for identifying new ligands for RBCC domain proteins.

© 2000 Academic Press

Keywords: transcriptional repression; KRAB domain; RBCC domain; RING finger; protein-protein interaction

*Corresponding author

Present addresses: G. E. Begg, Victor Chang Cardiac Research Institute, Darlinghurst, NSW 2010, Australia; D. E. Jensen, GlaxoWellcome, Inc. Research Triangle Park, North Carolina, 27709, USA.

Abbreviations used: KRAB, Krüppel associated box; RBCC, RING finger, B boxes, and coiled-coil region; Ni-NTA, Ni²⁺-nitrilo-tri-acetic acid; ND-PAGE, non-denaturing polyacrylamide gel electrophoresis; EMSA, electrophoretic mobility shift assay; EGS, ethylene glycol bis (succinimidylsuccinate); bv, baculovirus; HSF, heat shock transcription factor; ZFPs, zinc-finger proteins.

E-mail address for the corresponding author: rauscher@wista.wistar.upenn.edu

Introduction

The recurrent theme that has emerged from the structural study of proteins involved in signal transduction and transcriptional regulation is that multiple independently folded globular domains in these proteins often cooperate in macromolecular recognition. These domains are often recognizable by conserved signature amino acid sequence motifs and their organization within a protein often is the first clue about protein function. The bHLH and bZIP DNA-binding proteins have provided an important paradigm for protein multimerization and macromolecular recognition: each protein encodes adjacent, functionally separable dimerization and DNA-binding regions (Baxevanis & Vinson, 1993; Hurst, 1995). The dimerization function, which is mediated by a coiled-coil motif, is required to correctly orient the basic regions for interaction with DNA. With these principles in mind, we have been characterizing the RING-B box-coiled-coil (RBCC) family of nuclear regulatory proteins, which have been predicted to contain functionally separable multimerization and protein recognition domains (Borden, 1998).

The RBCC family of proteins contains the tripartite motif that most likely functions as a cooperative protein-protein interaction motif (Borden, 1998) (Figure 1(a) and (b)). The definitive element of the tripartite motif is the RING finger, which is a cysteine-rich motif of the form C3HC4 and binds two molecules of zinc in a unique cross-braced ligation system (Barlow *et al.*, 1994; Bellon *et al.*, 1997; Borden *et al.*, 1995a). The RING finger is found almost exclusively in the NH₂-terminal position in RBCC proteins and is likely to contribute either specificity and/or multimerization properties to the tripartite motif. More than 100 proteins encode for a RING finger motif, and mutational analyses have confirmed the requirement for the RING finger for proper biological function (for reviews, see Borden, 1998; Saurin *et al.*, 1996; Boddy *et al.*, 1997; Cao *et al.*, 1997). RING finger structures suggest that a common hydrophobic core is formed as a result of zinc-chelation and that the variable sequence between the conserved ligation residues provides specificity for protein recognition.

The second signature motif of the RBCC domain is the B-box, which is also a cysteine-rich zinc-binding motif of the form CHC3H2 (Reddy & Etkin, 1991) (Figure 1(b)). The B-box is an independently folded globular domain. Two B-box motifs are often found immediately carboxy-terminal to the RING finger in the RBCC domain. The NMR structure of the XNF7 B-box showed that only one zinc atom was bound with the other potential chelation residues unoccupied (Borden *et al.*, 1995b). XNF7 is required for embryonic development *via* a chromatin-binding function that requires the integrity of the B-box (Bellini *et al.*, 1995).

The third RBCC signature motif is a coiled-coil domain that is comprised of appropriately spaced hydrophobic residues predicted to form an

extended α -helical region (for a review, see Lupas, 1996) (Figure 1(b)). Multiple individual regions within the predicted coiled-coil region display amphipathic character. Moreover, many of the biological properties of RBCC proteins have been shown to be dependent on multimerization *via* this coiled-coil region (Saurin *et al.*, 1996).

Thus, RBCC proteins are defined by this tripartite motif, which is found in a spatially restricted organization in each protein. Clues to the function and/or potential ligands for RBCC proteins have come mainly from the studies of their biological properties. Most notably, RBCC proteins have been defined as potential proto-oncogenes: RFP, TIF1 α , and PML (Goddard *et al.*, 1991; Kakizuka *et al.*, 1991; Le Douarin *et al.*, 1995; Miki *et al.*, 1991; Takahashi *et al.*, 1985). The RBCC domain of each of these proteins is fused to heterologous proteins (as a result of chromosomal translocation) and results in the creation of oncogenes. The RBCC domain of PML protein is fused to the retinoic acid receptor alpha (RAR α) as a result of the t(15;17) translocation (Borden *et al.*, 1996; Goddard *et al.*, 1991; Kakizuka *et al.*, 1991), resulting in aberrant regulation of RAR target genes and a block in myelocytic differentiation (Grignani *et al.*, 1996). The PML-RAR α fusion homodimerizes and heterodimerizes with the wild-type PML and RXR: both of these activities require the intact coiled-coil region of the RBCC domain (Dyck *et al.*, 1994; Perez *et al.*, 1993; Weis *et al.*, 1994). Multimerization appears to sequester wild-type RAR/RXR receptors into discrete subnuclear domains termed ND10 or PODs (for a review, see Maul, 1998; Sternsdorf *et al.*, 1997). The wild-type PML protein functions as a growth suppressor *via* induction of apoptosis (Borden *et al.*, 1997). This activity requires both localization to ND10/PODs and cooperation between the RING, B-boxes and coiled-coil motifs. Thus, in this biological context, the PML-RBCC domain functions as a cooperative unit that presumably mediates specific protein-protein interactions.

The Krüppel associated box (KRAB) domain was originally identified as a conserved motif at the NH₂ terminus of zinc-finger proteins (ZFPs) (Bellefroid *et al.*, 1991) and was then shown to be a potent, DNA-binding-dependent transcriptional repression module (Margolin *et al.*, 1994; Pengue *et al.*, 1995; Vissing *et al.*, 1995; Witzgall *et al.*, 1994). The KRAB domain homology consists of approximately 75 amino acid residues and is predicted to fold into two amphipathic helices (Figure 1(c)). The minimal repression module is approximately 45 amino acid residues, and substitutions for conserved residues abolish repression (Margolin *et al.*, 1994). More than ten independently encoded KRAB domains have been demonstrated to be potent repressors, suggesting that this activity is a common property. Our laboratory purified and cloned KAP-1 as a protein that binds to the KRAB repression domain (Friedman *et al.*, 1996). Structurally, KAP-1 possesses all the

signature motifs for an emerging superfamily of transcriptional regulators, including TIF1 α and TIF1 γ (Friedman *et al.*, 1996; Le Douarin *et al.*, 1995; Venturini *et al.*, 1999) (Figure 1(a)). Further studies indicated that KAP-1 functions as a universal corepressor for KRAB domain proteins (Friedman *et al.*, 1996). KAP-1 is a 97 kDa nuclear phosphoprotein that possesses a canonical RBCC finger in the NH₂ terminus, and appears to be both necessary and sufficient for the KRAB domain binding. The carboxy terminus of KAP-1 includes a PHD finger and a bromodomain, and this region functions as a repressor when tethered to DNA using a heterologous DNA-binding domain, and mutations in either domain weaken this repression activity (unpublished results).

It has been firmly established that KAP-1 is required for KRAB domain-mediated repression. The evidence includes: (1) KAP-1 binds to multiple KRAB repression domains both *in vitro* and *in vivo*; (2) KRAB domain mutations that abolish repression decrease or eliminate KAP-1 binding; (3) overexpression of KAP-1 enhances KRAB-mediated repression; (4) the KRAB domain does not repress in cells that lack KAP-1. These results support a model in which KRAB-ZFPs bind a gene and repress transcription of the gene by recruiting the KAP-1 corepressor *via* the RBCC domain. Thus, the KRAB-KAP-1 interaction serves as a specific example from which to dissect molecular determinants for RBCC-ligand interactions. We have reconstituted the KRAB-KAP-1 complex using purified components and have employed a comprehensive set of biochemical approaches to analyze this interaction. Our data indicate that KAP-1-RBCC oligomerization is required for KRAB domain binding and that all three components of the tripartite motif cooperate in KRAB recognition.

Results

Direct interaction between the KAP-1-RBCC domain and the KOX1-KRAB domain

Previous studies using NH₂-terminal truncated KAP-1 proteins have suggested that the RING finger, B boxes, and coiled-coil structure are necessary for binding to the KRAB domain (Friedman *et al.*, 1996). However, all of these studies were performed using indirect protein-protein interaction assays either in complex mixtures *in vitro*, or in yeast two-hybrid assays. We sought to reconstitute the KRAB-KAP-1 interaction using purified recombinant components. The proteins were overexpressed using bacterial or baculovirus systems (see below) and purified using Ni-NTA chromatography to near-homogeneity (Figure 2(a)). The KAP-1-RBCC proteins produced in *Escherichia coli* were soluble and were purified under non-denaturing conditions. SDS-PAGE analysis (Figure 2(a)) revealed that the KAP-1-RBCC protein migrates with an apparent molecular mass of 46 kDa, a value that is almost identical with its predicted

molecular mass of 45.9 kDa. The KAP-1-RBCC mutants show similar degrees of purity and migrate identically in SDS-PAGE (Figure 2(a)).

The KRAB domain that we have utilized in these studies corresponds to amino acid residues 1-90 of the KOX1 zinc finger protein (Figure 1(c)). Although the KRAB domain is found in more than 100 independent proteins, we selected the KRAB domain from KOX1 for the following reasons: (1) the KOX1-KRAB domain was originally utilized to isolate the KAP-1 corepressor, and mutations in this domain that concomitantly abolish repression and KAP-1 binding are well characterized (Friedman *et al.*, 1996). (2) The KOX1-KRAB domain is highly expressed in *E. coli* autonomously or as a GAL4 fusion and is well behaved in protein reconstitution assays. (3) The GAL4-KOX1-KRAB 1-90 fusion (designated GAL4-KRAB) is a potent, KAP-1-dependent and DNA binding-dependent transcriptional repressor *in vivo* (Margolin *et al.*, 1994). The wild-type GAL4-KRAB and a mutant version (DV18,19AA, which lacks repression activity) were expressed and purified from *E. coli* (Figure 2(a)). To detect whether the purified KAP-1-RBCC protein was able to form a stable complex with a DNA-bound KRAB domain, we used the electrophoretic mobility shift-supershift assay (EMSA). Binding of the GAL4-KRAB protein to a canonical ³²P-labeled synthetic oligonucleotide containing the GAL4 recognition sequence yielded an expected mobility shift (Figure 2(b), lane 1). When increasing amounts of purified KAP-1-RBCC protein were incubated with the GAL4-KRAB protein and DNA, a new mobility shift was observed (Figure 2(b), lanes 2-6). This supershift contains the ternary complex of DNA:GAL4-KRAB:KAP-1-RBCC, as characterized by antibody inhibition/supershift and competition with unlabeled competitor DNA (data not shown). Formation of this complex requires the GAL4-KRAB DNA binding subunit since no shift is seen with KAP-1-RBCC alone (data not shown). The GAL4-KRAB(DV-AA) protein, which shows severely impaired repression function *in vivo* readily bound to ³²P-labeled GAL4 DNA but failed to generate the ternary complex in the presence of increasing amounts of KAP-1-RBCC protein (Figure 2(b), lanes 7-12), in agreement with previous reports (Friedman *et al.*, 1996). These data strongly suggest that the KAP-1-KRAB interaction is direct and highly specific, and that the RBCC region is necessary and sufficient for this interaction.

The RING finger, B boxes, and the coiled-coil region are critical for the formation of the KRAB:KAP-1-RBCC complex

To define the sub-domains of the KAP-1-RBCC domain integral to its interaction with the KRAB domain, we made independent mutations in the RING finger, B boxes, and the coiled-coil region (Figure 1(d)). Mutant proteins were expressed and purified in *E. coli* (Figure 2(a)), and analyzed for

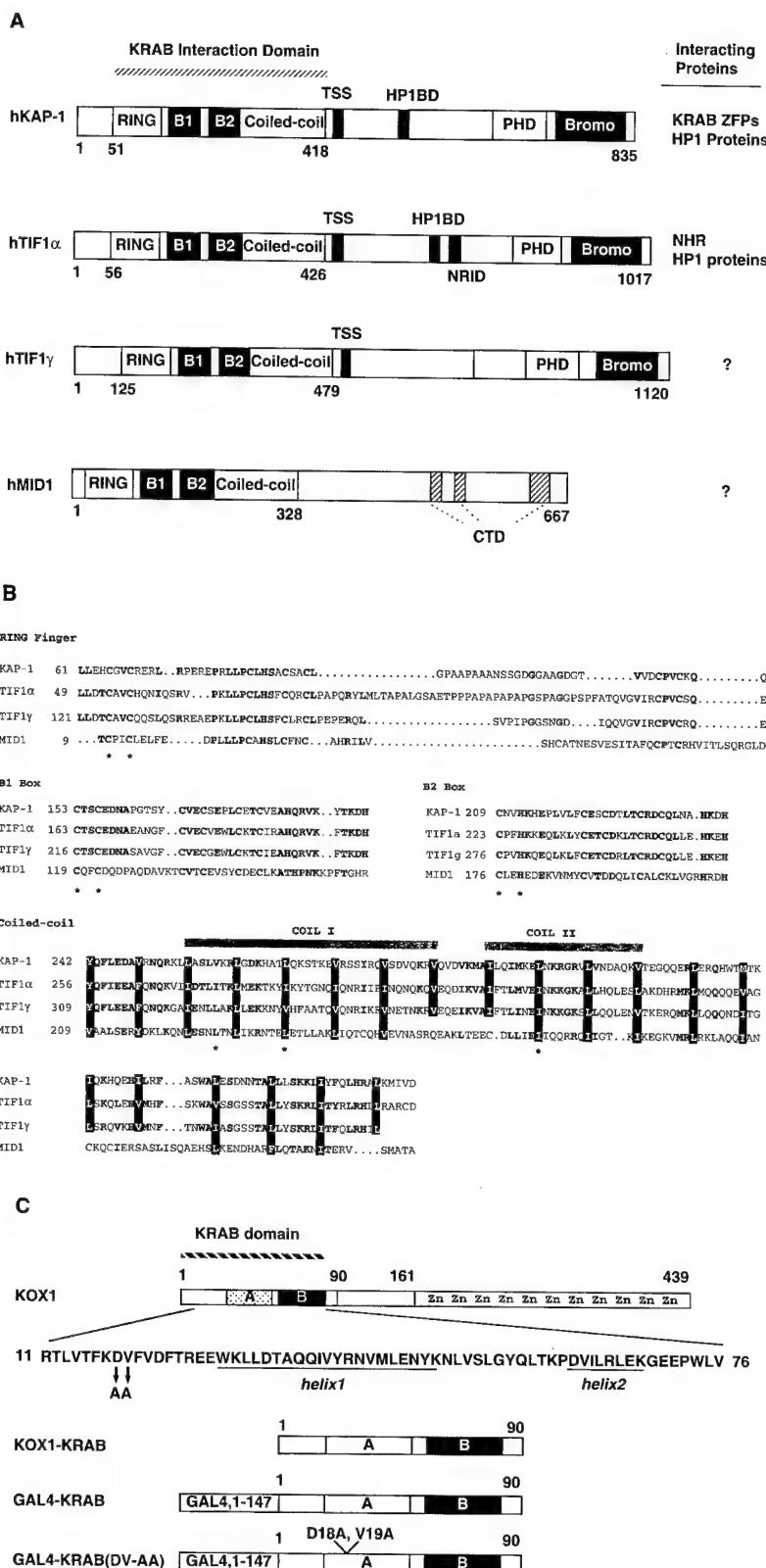
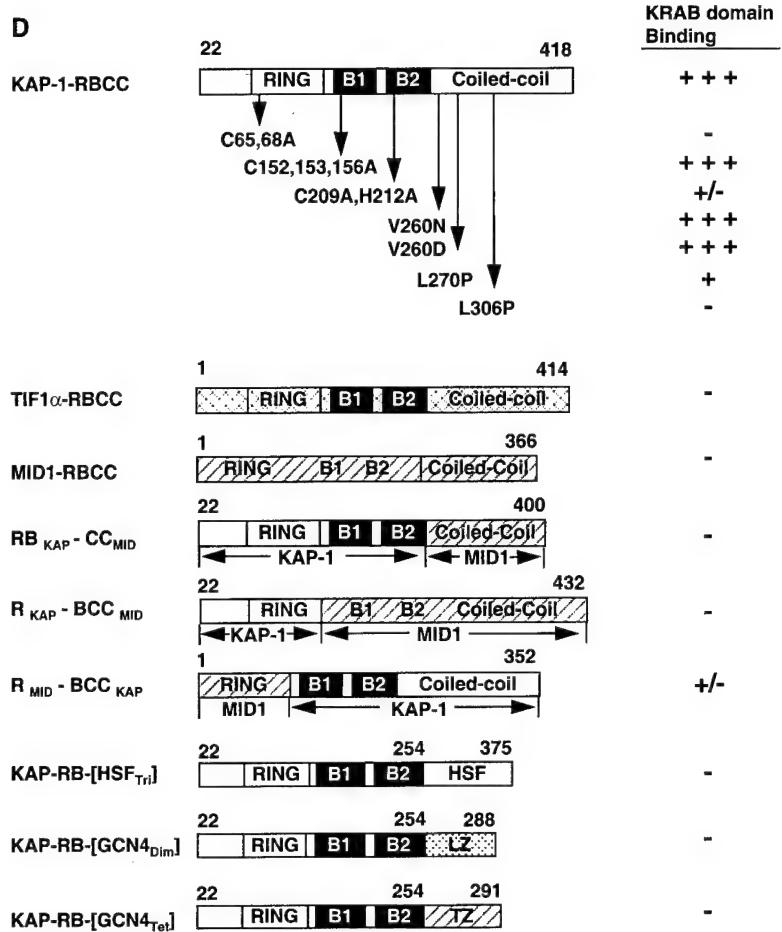


Figure 1 (legend opposite)



coiled-coil domain, hydrophobic amino acid residues defining the heptad repeats of each putative coil are highlighted in black. The relative position of the first and the second predicted coils of the coiled-coil region are indicated by the gray bars above the sequence alignment. The numbers at the left refer to the amino acid positions in the corresponding proteins. Database accession numbers: hKAP-1 (998813); hTIF1 α (AF119042); hTIF1 γ (AF119043); hMID1 (Y13667). (c) A representation of the KOX1-ZFP and recombinant derivatives used in this study. The amino acid sequence of the minimal KRAB domain of KOX1 is illustrated. The position of the DV18,19AA mutation, which disrupts KRAB-mediated repression and association with KAP-1, is indicated (Margolin *et al.*, 1994; Friedman *et al.*, 1996). The KRAB domain of KOX1, residues 1-90, was expressed without a DNA-binding domain and as a heterologous fusion protein with the DNA-binding domain of GAL4, residues 1-147. (d) A diagram of the KAP-1-RBCC domain (wild-type and mutants) and organization of heterologous KAP-1-RBCC fusion proteins. The KAP-1-RBCC (CC65,68AA) is a double substitution in the RING finger. The KAP-1-RBCC (CCC152,153,156AAA) is a triple substitution in the B1 box. The KAP-1-RBCC (C209A,H212A) is a double substitution in the B2 box. The KAP-1-RBCC (V260N), KAP-1-RBCC (V260D), and KAP-1-RBCC (L270P) are single substitutions in the first predicted coil of the coiled-coil region. The KAP-1-RBCC (L306P) is a single substitution in the second predicted coil of the coiled-coil region. The TIF1 α -RBCC domain consists of amino acid residues 1-414 of the TIF1 α protein. The MID1-RBCC domain consists of residues 1-366 of the MID1 protein. The RB_{KAP}-CC_{MID} is a fusion between the RING finger, B boxes of KAP-1 (residues 22-254) and the coiled-coil region of MID1 (residues 220-366). The R_{KAP}-BCC_{MID} is a fusion between the RING finger of KAP-1 (residues 22-138) and the B boxes, and coiled-coil region of MID1 (residues 72-366). The R_{MID}-BCC_{KAP} is a fusion between the RING finger of MID1 (residues 1-71) and the B boxes, and coiled-coil region of KAP-1 (residues 139-418). The KAP-RB-[HSF_{Tri}], KAP-RB-[GCN4_{Dim}], and KAP-RB-[GCN4_{Tet}] are heterologous fusions between residues 22-254 of KAP-1 and the trimerization domain of HSF (residues 321-441), the leucine zipper of GCN4 (residues 253-281), and the synthetic tetramerization domain of GCN4 (residues 250-281). A summary of the KRAB binding potential is indicated by plus signs at the right.

KRAB domain binding (Figure 3). Our initial mutational strategy was to target the critical metal chelation residues. For the RING finger and B-boxes, we mutated the potential Zn chelation residues in

Figure 1. A diagram illustrating the architecture of the KAP-1/TIF1 family of transcriptional regulatory proteins and the KOX1/KRAB-ZFP protein. (a) A representation of proteins containing the tripartite RBCC motif, including hKAP-1, hTIF1 α , hTIF1 γ , and hMID1. The numbers represent amino acid positions. RING, RING finger; B1 and B2, B-box; TSS, TIF1 signature sequence (Venturini *et al.*, 1999); HP1BD, HP1 binding domain (Ryan *et al.*, 1999); PHD, plant homeodomain; Bromo, bromo domain; NRID, nuclear receptor interaction domain; CTD, carboxy-terminal domain. The column to the right summarizes known protein-protein interactions: NHR, nuclear hormone receptor; HP1, heterochromatin protein-1. (b) Amino acid alignment of the RING finger, B boxes, and coiled-coil region of the hKAP-1, hTIF1 α , hTIF1 γ , and hMID1 proteins. The alignments were generated using the CLUSTAL program. The amino acid residues that are identical and/or conserved in all four proteins are in bold. The asterisks below indicate the amino acids targeted for mutational analysis in this study. The periods represent spaces introduced to obtain maximal alignment. The spacing between the B2-box and the first coil of KAP-1 is 15 amino acid residues. Within the

pairs (or multiples) to lessen the possibility that alternative chelation residues from adjacent sequence could be recruited to restore metal binding. A double substitution (CC65,68AA) was made

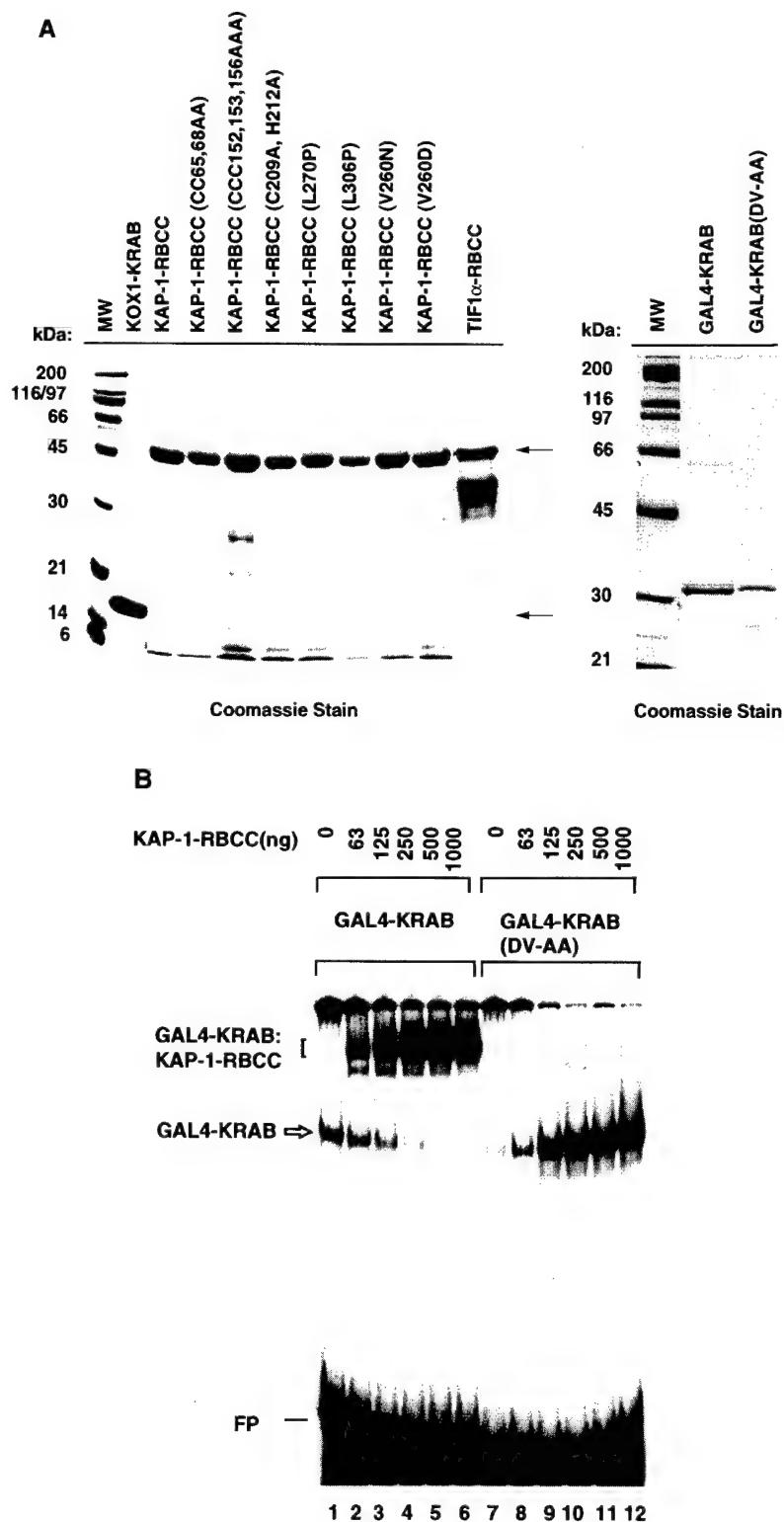


Figure 2. Purification, and functional analysis of recombinant proteins expressed in *E. coli*. (a) A Coomassie blue-stained SDS/polyacrylamide gel of the purified recombinant KOX1-KRAB, KAP-1-RBCC, and GAL4-KRAB fusion proteins. From 3 to 5 μ g of the purified KOX1-KRAB, KAP-1-RBCC, and TIF1 α -RBCC proteins was loaded per a lane. 1 μ g of GAL4-KRAB (wild-type and DV-AA mutant) protein was loaded per lane. The arrow indicates the recombinant KOX1-KRAB protein. The arrowhead indicates the various KAP-1-RBCC and TIF1 α -RBCC proteins. (b) Binding of purified bacterial KAP-1-RBCC to a wild-type KRAB domain but not a mutant form as detected by EMSA. EMSA was performed as described in Materials and Methods. Each reaction contained a constant amount of purified GAL4-KRAB (100 ng), or GAL4-KRAB (DV-AA) (100 ng) (a), a 32 P-labeled DNA probe containing the canonical GAL4 binding site, and increasing amounts of purified bacterial KAP-1-RBCC (a). The DNA:protein complexes were resolved in non-denaturing polyacrylamide gels. The open arrow indicates the DNA:GAL4-KRAB binary complex. The slower migrating DNA:GAL4-KRAB:KAP-1-RBCC complex is represented by the bracket. FP, free probe.

at two conserved cysteine residues in the consensus C3HC4 signature (Figure 1(b)) that have been shown to be critical for both coordination of zinc and for stabilization of the hydrophobic core

(Saurin *et al.*, 1996). Mutations in the RING finger completely abolished the ability of the KAP-1-RBCC domain to interact with the KRAB domain (Figures 1(d) and 3(a) and (c)).

The second structural element in the RBCC tripartite motif is the B box, which is characterized by a signature spacing of cysteine and histidine residues (Borden, 1998) (Figure 1(b)). Surprisingly, a triple substitution (CCC 152,153,156 AAA) in the NH₂-terminal B box had no effect on the KAP-1-RBCC binding to the KRAB domain (Figures 1(d) and 3(c)). However, a double substitution (C209A,H212A) in the second B box (Figure 1(b)) strongly reduced the KAP-1-RBCC and KRAB domain interaction (Figures 1(d) and 3(c)).

The coiled-coil domain of KAP-1 is predicted to fold into two potential leucine zipper-like motifs (Wolf *et al.*, 1997) (Figures 1(b) and 3(b)). Each region is likely to be an amphipathic α -helix, having a seven-residue repeat (**a.b.c.d.e.f.g**)_n, with hydrophobic residues at positions **a** and **d** and polar residues at the other positions (Figure 3(b)). We altered the leucine zipper motif by substituting a proline residue for a hydrophobic amino acid, which would be predicted to disrupt the α -helix structure. The L270P mutation at position **d** in coil I shows an intermediate effect on the KRAB:KAP-1 interaction (Figures 1(d) and 3(a) and (c)). The L306P mutation at position **d** in coil II was found to completely disrupt the KRAB:KAP-1 interaction (Figures 1(d) and 3(c)). We also tested whether the substitution of a polar amino acid for a hydrophobic amino acid at position **a** would interfere with the KRAB:KAP-1 interaction (Figure 3(b)). However, change of Val260 to either Asn or Asp displayed little effect on the KRAB:KAP-1-RBCC interaction (Figures 1(d) and 3(c)). Together, the above data suggest that an intact RING finger, B boxes, and the coiled-coil domain are each important for the stable binding of the KAP-1-RBCC to a KRAB domain.

Since TIF1 α is 40% identical with KAP-1 within the RBCC domain (Figure 1(b)), we investigated whether the RBCC domain of TIF1 α could interact with the KRAB domain. We expressed and purified the RBCC domain of TIF1 α (residues 1-414) (Figure 2(a)). The RBCC domain of TIF1 α displayed no binding to the KRAB domain (Figure 3(c)), an observation that is in contrast with a previous report (Moosmann *et al.*, 1996). It is possible that the KRAB-TIF1 α interaction reported by Moosmann was indirect, as it was detected only in yeast two-hybrid assays and was independent of the RING finger and B boxes.

KAP-1-RBCC exists in an oligomeric state

To investigate the mechanisms behind the KRAB:KAP-1 interaction, we employed biochemical and biophysical analyses of the KRAB domain, the KAP-1-RBCC domain, and the complex. Gel-filtration chromatography was used to estimate the native hydrodynamic size (Stokes radius) of the KAP-1-RBCC protein, which eluted with an apparent molecular mass of 158 kDa (Figure 4(a)-I). This value is roughly three times the molecular mass of the KAP-1-RBCC monomer, suggesting that the

KAP-1-RBCC domain is an apparent trimer. This result is consistent with multi-coil scoring analysis (Wolf *et al.*, 1997) for the KAP-1 coiled-coil region.

To determine the oligomeric state of the KAP-1-RBCC in the presence of the native KRAB domain, a KRAB:KAP-1-RBCC complex was pre-formed in solution, and then subjected to gel filtration. This experiment required purified KOX1-KRAB protein from *E. coli*, which initially remained profoundly insoluble. We also attempted to express the KRAB domains from ZNF133, ZNF140, ZNF141, KRK-1, and EEK-1, which also proved insoluble. We finally succeeded in producing soluble KOX1-KRAB protein by employing a protocol that refolds the protein while bound to the Ni-NTA resin (Shi *et al.*, 1997). The KOX1-KRAB domain was purified to greater than 98% purity and migrated on SDS-PAGE with an apparent molecular mass of 15.5 kDa, which was slightly larger than the predicted molecular mass of 13.9 kDa (Figure 2(a)). Full analysis of this KOX1-KRAB protein will be described elsewhere. However, secondary structure analysis using circular dichroism spectroscopy indicated that the protein was properly folded, and the protein was highly active in binding assays. As shown in Figure 4(a)-III, the KAP-1-RBCC domain in KOX1-KRAB:KAP-1-RBCC complex eluted from the gel-filtration column as a distribution of species, most likely existing in a complex equilibrium, where the majority eluted with an apparent molecular mass consistent with a trimer. Note the slightly larger mass observed for the KAP-1-RBCC complexed with the KRAB domain (fraction 11, Figure 4(a)-III) compared to uncomplexed KAP-1-RBCC (fraction 12, Figure 4(a)-I). However, other species exist and these can co-elute with the KRAB domain. The majority of the KOX1-KRAB protein eluted at 670 kDa, suggesting that this protein aggregates when it is not complexed with KAP-1-RBCC (Figure 4(a)-II). To confirm that this complex was stable during chromatography, the peak fraction (11) containing the complex was re-chromatographed under identical conditions. The peak of the KOX1-KRAB:KAP-1-RBCC complex eluted exactly as the previous run, with an apparent molecular mass slightly larger than 158 kDa (Figure 4(a)-IV). We conclude that (1) the KOX1-KRAB:KAP-1-RBCC complex can be readily formed in solution and is stable to gel filtration under these conditions; (2) there is a small but significant increase in the Stokes radius of the complex compared with KAP-1-RBCC alone; and (3) the KAP-1-RBCC domain exists as an oligomer, probably a trimer in the absence or presence of the KRAB domain.

We next used velocity sedimentation in sucrose gradients to estimate the native mass of the KOX1-KRAB:KAP-1-RBCC complex and its components. The KAP-1-RBCC protein alone sedimented with an apparent mass of 66 kDa (Figure 4(b)-I), suggesting either that the oligomeric state of the KAP-1-RBCC in solution was not stable to the conditions of sucrose gradient sedimentation, or that

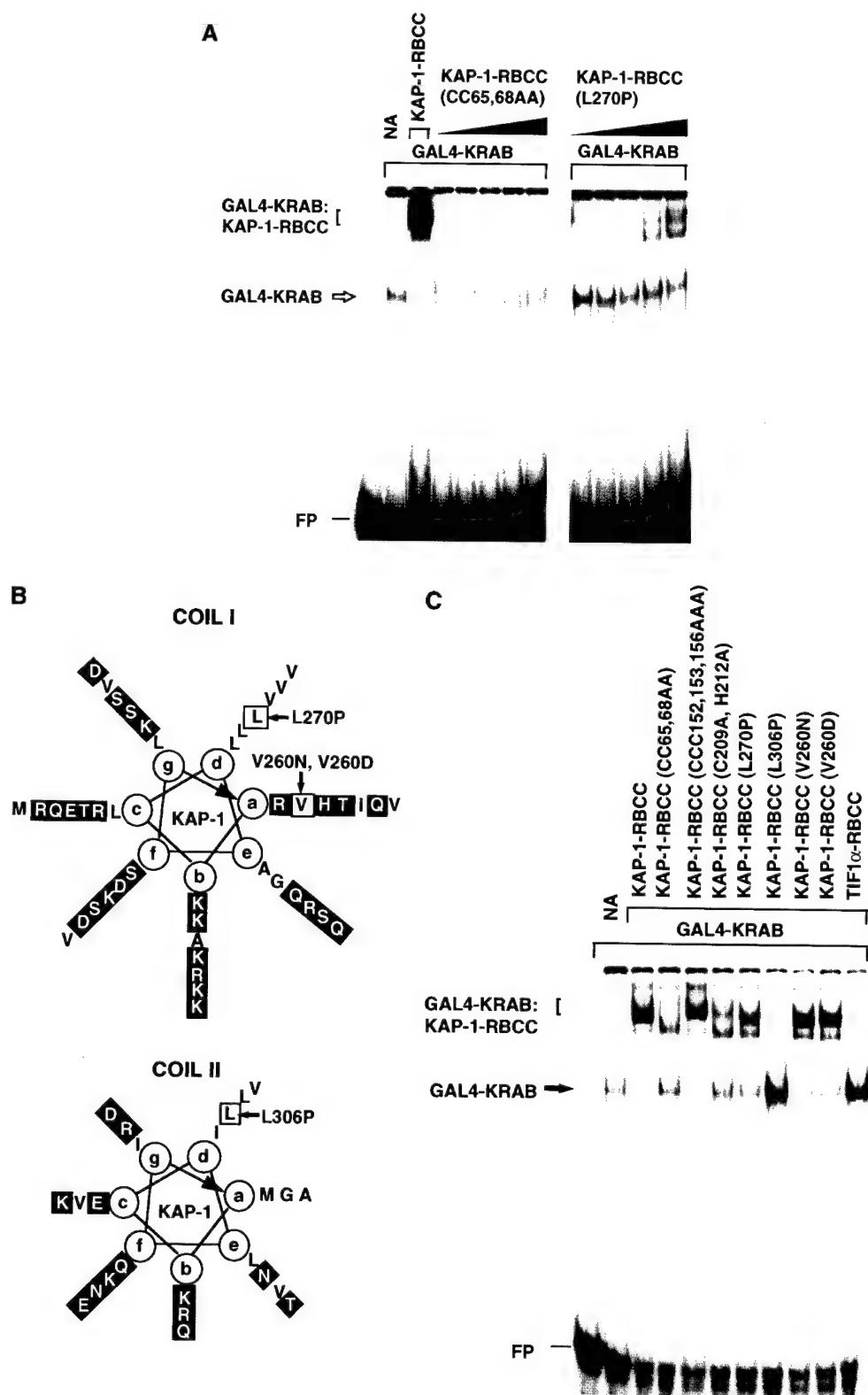


Figure 3 (legend opposite)

the protein was in an extended configuration that results in the relatively large Stokes radius observed by gel filtration. The KOX1-KRAB protein alone sedimented with apparent masses of 670 and 13.7 kDa (Figure 4(b)-II). However, pre-formation of a KOX1-KRAB:KAP-1-RBCC complex prior to centrifugation resulted in a dramatic shift in the sedimentation of KAP-1-RBCC domain to an apparent mass of approximately 158 kDa (Figure 4(b)-III). The KOX1-KRAB domain was detected with apparent molecular masses of 670, 158, and 25 kDa, most likely representing aggregated, KAP-1-RBCC-complexed, and monomeric or dimeric uncomplexed KOX1-KRAB protein, respectively. This sedimentation analysis supports the conclusion that the KAP-1-RBCC is oligomeric and that the KRAB domain may stabilize the oligomerization of the KAP-1-RBCC under these sucrose gradient conditions.

We next utilized non-denaturing-PAGE (ND-PAGE) to estimate the native size of KAP-1-RBCC. We calculated the mobilities of the protein standards and of KAP-1-RBCC relative to the tracking dye in a variety of polyacrylamide gel concentrations and used this to generate plots of mobility *versus* polyacrylamide concentration. The slopes of such plots represent the "retardation coefficients" of the proteins (Bryan, 1977). A log-log plot was generated in which the negative retardation coefficients of the standards were plotted against the known native molecular masses of the standards (Figure 4(c)). From this calibration curve, we estimated that the KAP-1-RBCC has a native size of 200 kDa under this condition. These data further substantiate the suggestion that the KAP-1-RBCC protein exists as an oligomer.

To further investigate the oligomeric state of the KAP-1-RBCC domain, sedimentation equilibrium experiments were performed using analytical ultracentrifugation. For each experiment, analyses were performed at 4°C and 25°C, and the concentration of protein *versus* radius data was fitted with various models of self-association, using non-linear regression (Johnson *et al.*, 1981). The KAP-1-RBCC protein showed non-ideal behavior under these

conditions. For two independent experiments, each containing three samples of differing KAP-1-RBCC protein concentration, the data were best described by a model containing monomer, trimer, and hexamer terms (Figure 5). However, when multiple data sets were fitted simultaneously, neither the monomer-trimer nor the monomer-hexamer association could be described by a single equilibrium constant, indicating that these associations are not reversible on the time-scale of the experiment. The apparent equilibrium constants describing the monomer-trimer and the monomer-hexamer association increased with decreasing protein concentration loaded, which suggests that this apparent irreversibility is due to the presence of incompetent trimer and hexamers, i.e. oligomers that do not dissociate. There may also be a reversible association, however, as there was some redistribution of the species (most obviously a reduction in the proportion of hexamer) when the temperature was raised from 4°C to 25°C in a single experiment. Unfortunately, we were unable to obtain useful data starting with a pre-formed KOX1-KRAB:KAP-1-RBCC complex, because the protein concentrations of the complex that could be obtained were not high enough and a significant portion of the sample formed aggregates under the conditions required for the ultracentrifugation experiments. Nonetheless, these analyses suggest that KAP-1-RBCC is predominantly a trimer in solution and a trimer-hexamer equilibrium may play a physiological role.

We next employed chemical cross-linking analyses with ethylene glycol bis (succinimidylsuccinate) (EGS) to determine the oligomeric state of the KAP-1-RBCC proteins (wild-type and mutants) (Figure 6). Treatment of purified KAP-1-RBCC protein with low concentrations of EGS yielded four distinct cross-linked species with apparent molecular masses of 100, 135, 210, and 270–300 kDa, potentially corresponding to dimeric, trimeric, tetrameric, and hexameric forms of the KAP-1-RBCC domain (Figure 6(a)). The monomeric form of the KAP-1-RBCC domain migrated at 49 kDa in SDS-PAGE. Treatment of a pre-formed KOX1-KRAB:

Figure 3. Mutations within the RING finger, B boxes, and the coiled-coil region of the KAP-1-RBCC abolish interaction with the KRAB domain *in vitro*. (a) Binding of the wild-type KAP-1-RBCC or the mutant KAP-1-RBCC (CC65,68AA)/(L270P) proteins to the wild-type GAL4-KRAB, as detected by EMSA. Wild-type KAP-1-RBCC (1 µg) or increasing amounts of mutant protein (63, 125, 250, 500, 1000 ng) (Figure 2(a)) were incubated with the wild-type GAL4-KRAB protein. The open arrow indicates DNA:GAL4-KRAB mobility shift. The bracket indicates the DNA:GAL4-KRAB:KAP-1-RBCC ternary complex. (b) An illustration of the putative heptad repeat for coil I (residues 253–294) and coil II (residues 299–322) of the KAP-1 coiled-coil region. Note the potential amphipathic nature of the putative helices. The heptad positions are labeled by letters **a** through **g** in the helix, where hydrophobic residues at positions **a** and **d** might constitute the cores of homo-trimeric oligomerization domain. The hydrophilic residues at the other positions are indicated with the black background. The two leucine residues (L270 and L306) mutated to proline, and the valine residue (V260) mutated to asparagine or aspartic acid are boxed. (c) The interaction between the KRAB domain and the KAP-1-RBCC domain is dependent upon the integrity of the RBCC domain. The binding of GAL4-KRAB to KAP-1-RBCC or various mutant KAP-1-RBCC proteins, as detected by EMSA. A constant amount of GAL4-KRAB (100 ng), 1 µg of each KAP-1-RBCC protein (wild-type or mutant), and 1 µg of TIF1 α -RBCC protein was added to each reaction. The arrow indicates the mobility shift for the GAL4-KRAB protein. The bracket indicates the DNA:GAL4-KRAB:KAP-1-RBCC ternary complex. FP, free probe.

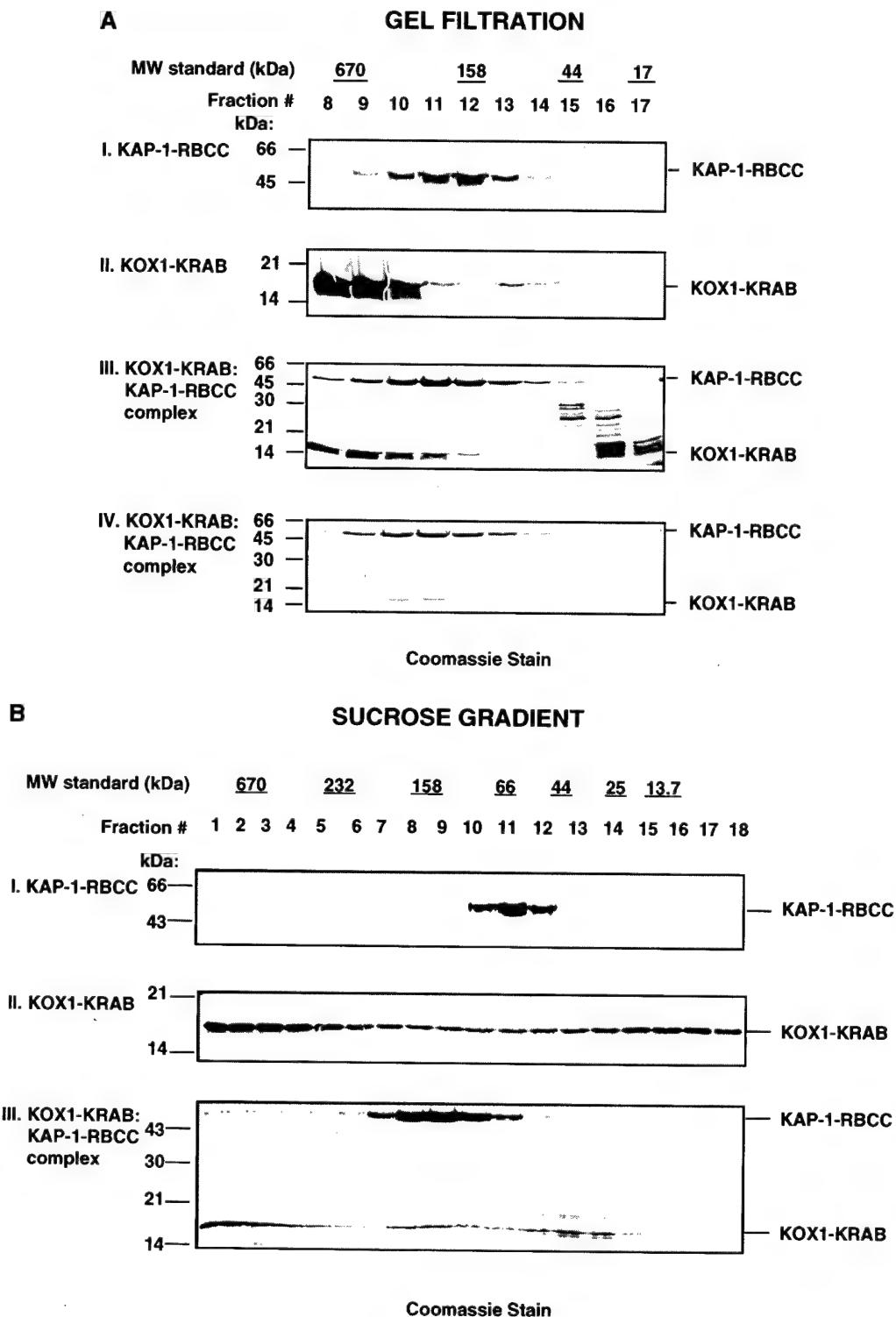


Figure 4 (legend opposite)

KAP-1-RBCC complex with low concentration of EGS yielded the same set of cross-linked species as was observed for the KAP-1-RBCC domain. However, a cross-linked species consistent with the for-

mation of a trimeric KAP-1-RBCC domain appeared as the predominant species even at very low levels of EGS (Figure 6(b)). These data suggest that the KRAB domain preferentially binds to a

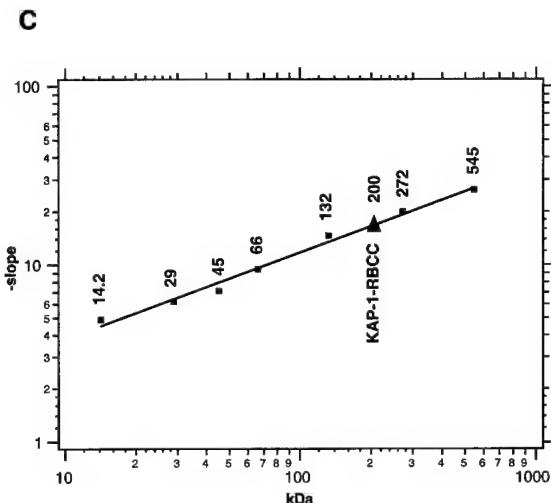


Figure 4. Analyses of the KAP-1-RBCC motif, the KOX1-KRAB domain, and a KOX1-KRAB:KAP-1-RBCC complex by gel-filtration, sucrose gradient sedimentation, and ND-PAGE. (a) Gel-filtration analysis of KAP-1-RBCC protein, KOX1-KRAB protein, and a pre-formed KOX1-KRAB:KAP-1-RBCC complex. The proteins were resolved on a Superdex 200 column. The fractions were collected and proteins were analyzed by SDS-PAGE. The fraction numbers are labeled at the top of the Figure. The migration of protein standards is indicated above the gels. The protein standards were run in a parallel gel-filtration experiment under identical conditions. (I) Elution profile of native KAP-1-RBCC protein; (II) the elution profile of KOX1-KRAB protein; (III) the elution profile for a pre-formed KOX1-KRAB:KAP-1-RBCC complex; (IV) the profile of the material from fraction 11 (III) re-chromatographed on the Gel-filtration column. (b) Sucrose gradient sedimentation analysis of the KAP-1-RBCC domain, KOX1-KRAB domain, and the KOX1-KRAB:KAP-1-RBCC complex. The proteins from sucrose gradient fractions were analyzed by SDS-PAGE. The fraction numbers are indicated at the top of each lane. The peak position of protein standards obtained from gradients run in parallel is indicated above the gels. (c) Determination of the molecular mass of native KAP-1-RBCC by ND-PAGE. The calibration curve prepared from Ferguson plots is shown (see Materials and Methods for details). The known masses of standards and the calculated mass of native KAP-1-RBCC are given (in kDa) above the curve.

discrete KAP-1-RBCC oligomer and stabilizes this oligomeric form, which is most consistent with a trimer. These cross-linking results are consistent with the sedimentation observed for a pre-formed KOX1-KRAB:KAP-1-RBCC complex in sucrose gradients (Figure 4(b)-III).

Analysis of the mutant KAP-1-RBCC proteins (CC65,68AA) and (L270P), which were severely impaired in KRAB binding (Figures 1(d) and 3(a) and (c)), showed little to no cross-linked species (dimer, trimer, or hexamer) under the same cross-linking conditions (Figure 6(c) and (d)). Furthermore, these mutant proteins migrated with an apparent molecular mass of approximately 66 kDa in sucrose gradients and this migration was not altered by pre-incubation with the KOX1-KRAB protein (data not shown). These data suggest that the formation of the KOX1-KRAB:KAP-1-RBCC complex likely requires the proper oligomeric state of the KAP-1-RBCC, which is apparently a trimer, and that KRAB binding stabilizes this state in solution.

The studies described above were exclusively performed with bacterial-expressed proteins that lack any eukaryotic post-translational modifications and the potential to interact with endogenous cellular partner proteins. To determine if our observations extend to eukaryotic cell-expressed

proteins, we overexpressed the KAP-1-RBCC and its mutant derivatives in Sf9 insect cells using a baculovirus expression vector. Full-length wild-type KAP-1, KAP-1-RBCC and mutants thereof were highly expressed as soluble proteins in Sf9 cells. The bv.KAP-1-RBCC was purified under native conditions and migrated as a single polypeptide band consistent with its predicted monomeric molecular mass of 47.7 kDa (data not shown). This behavior suggests that bv.KAP-1-RBCC is not significantly subjected to extensive post-translational modification, as detected by SDS-PAGE when expressed in eukaryotic cells. The bv.KAP-1-RBCC protein behaved identically with the *E. coli*-expressed protein in all of the biochemical binding assays and was equally active in KRAB domain interaction (data not shown).

To estimate the native mass and oligomeric states of KAP-1 proteins (wild-type and mutants) expressed in the baculovirus expression system, lysates from infected cells that contain the expressed proteins were subjected to gel-filtration chromatography and detected in a Western blot assay (Figure 7(a)). The full-length KAP-1 protein eluted with an apparent molecular mass of 500 to 670 kDa (Figure 7(a)-I). This elution profile was identical with both bacterial-expressed full-length KAP-1 and endogenous KAP-1 derived from

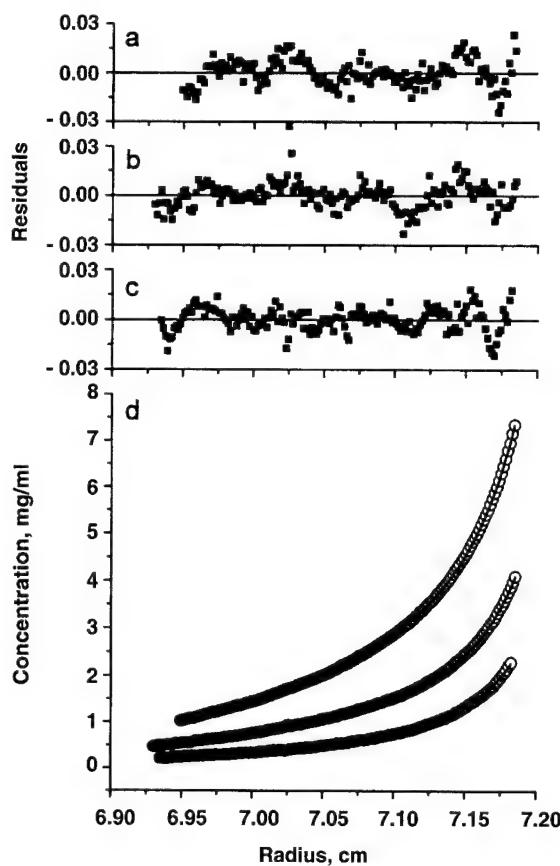


Figure 5. Representation of analytical ultracentrifugation analysis of KAP-1-RBCC. Sedimentation equilibrium analysis of KAP-1-RBCC was performed with three different loading concentrations in separate cells at 13,000 rpm and 4°C. The circles in the main panel (d) show the concentration *versus* radius data for the three cells at equilibrium. The three data sets were fitted separately with a model describing monomers, trimers, and hexamers, using the non-linear regression program NONLIN (Johnson *et al.*, 1981). The calculated fit is represented by the continuous lines. (a)-(c) show the residuals for the data points of the fitted curves at the three protein concentrations displayed from highest concentration (top) to lowest concentration.

mammalian cell nuclear extracts (unpublished results). Since the apparent monomeric molecular mass measured by SDS-PAGE of full-length KAP-1 is 97 kDa, this result suggests that native full-length KAP-1 was an oligomer, possibly as large as a hexamer. The bv.KAP-1-RBCC protein eluted with an identical Stokes radius as the *E. coli* KAP-1-RBCC protein at approximately 158 kDa (Figure 7(a)-II), suggesting that *E. coli* and bv-expressed KAP-1-RBCC share a common oligomeric state. Mutations in the RING finger (CC65,68AA) and the coiled-coil region (L270P) of the KAP-1-RBCC appeared to alter oligomerization properties, as these mutant proteins eluted with an apparent molecular mass of 100 kDa (Figure 7(a)-

III,IV), and also failed to bind GAL4-KRAB in the EMSA assay (Figure 8(a), lanes 3 and 4). These data are consistent with those observed for proteins produced in *E. coli*, and strengthen the argument that the KAP-1-RBCC must properly oligomerize for KRAB binding, which is independent of post-translational modifications.

Formation and detection of KAP-1:KAP-1-RBCC hetero-oligomers *in vivo*

To reconstitute hetero-oligomers of the KAP-1-RBCC domain with full-length KAP-1 *in vivo*, we co-expressed full-length KAP-1 with either wild-type or mutant KAP-1-RBCC proteins *via* co-infection of Sf9 cell cultures (Figure 7(b)). Because the KAP-1-RBCC encodes an NH₂-terminal 6His tag, its association with wild-type full-length KAP-1 could be detected by Ni-NTA chromatography, followed by Western blot analysis using an anti-KAP-1 serum that detects amino acid residues 423-589, a region that is not present in the KAP-1-RBCC protein (Friedman *et al.*, 1996) (Figure 7(b)). The KAP-1-RBCC bound to the Ni-NTA resin was detected with an anti-His monoclonal antibody. The full-length bv.KAP-1 (which has no 6His tag) did not bind to Ni-NTA resin (Figure 7(b), lane 1). When the Ni-NTA resin was incubated with Sf9 cell lysates, in which full-length KAP-1 protein was co-expressed with the KAP-1-RBCC, we observed retention of the full-length KAP-1 protein (Figure 7(b), lane 2). These results suggest that the KAP-1-RBCC domain can hetero-oligomerize with the full-length KAP-1 protein. Interestingly, simple mixing of cell lysates that contained independently expressed proteins failed to recapitulate this interaction (data not shown). Thus, the formation of hetero-oligomers between full-length KAP-1 and the KAP-1-RBCC domain may be coupled to translation or other post-translation processes, an observation that has been made for other oligomeric coiled-coil proteins. The KAP-1-RBCC mutants CC65,68AA and L270P also efficiently associated with full-length KAP-1 *in vivo* using the co-infection assay (Figure 7(b), lanes 3 and 4). This result suggests that these mutations inhibit the homo-oligomerization of the RBCC domain but that these RBCC mutants can still hetero-oligomerize with a wild-type RBCC domain *in vivo*.

Since the CC65,68AA and L270P mutant RBCC proteins could apparently hetero-oligomerize with full-length KAP-1, based upon the data above, we tested the ability of these hetero-oligomers to interact with the KRAB domain in the EMSA assay. The bv.KAP-1-RBCC formed the characteristic complex with the DNA-bound GAL4-KRAB protein (Figure 8(a), lane 2). The mutant proteins CC65,68AA and L270P each failed to form this ternary complex (Figure 8(a), lanes 3 and 4), a result that is consistent with our previous findings, in which *E. coli*-derived mutant RBCC proteins failed to bind the KRAB domain (Figures 1(d) and 3(a) and (c)). The full-length bv.KAP-1 protein

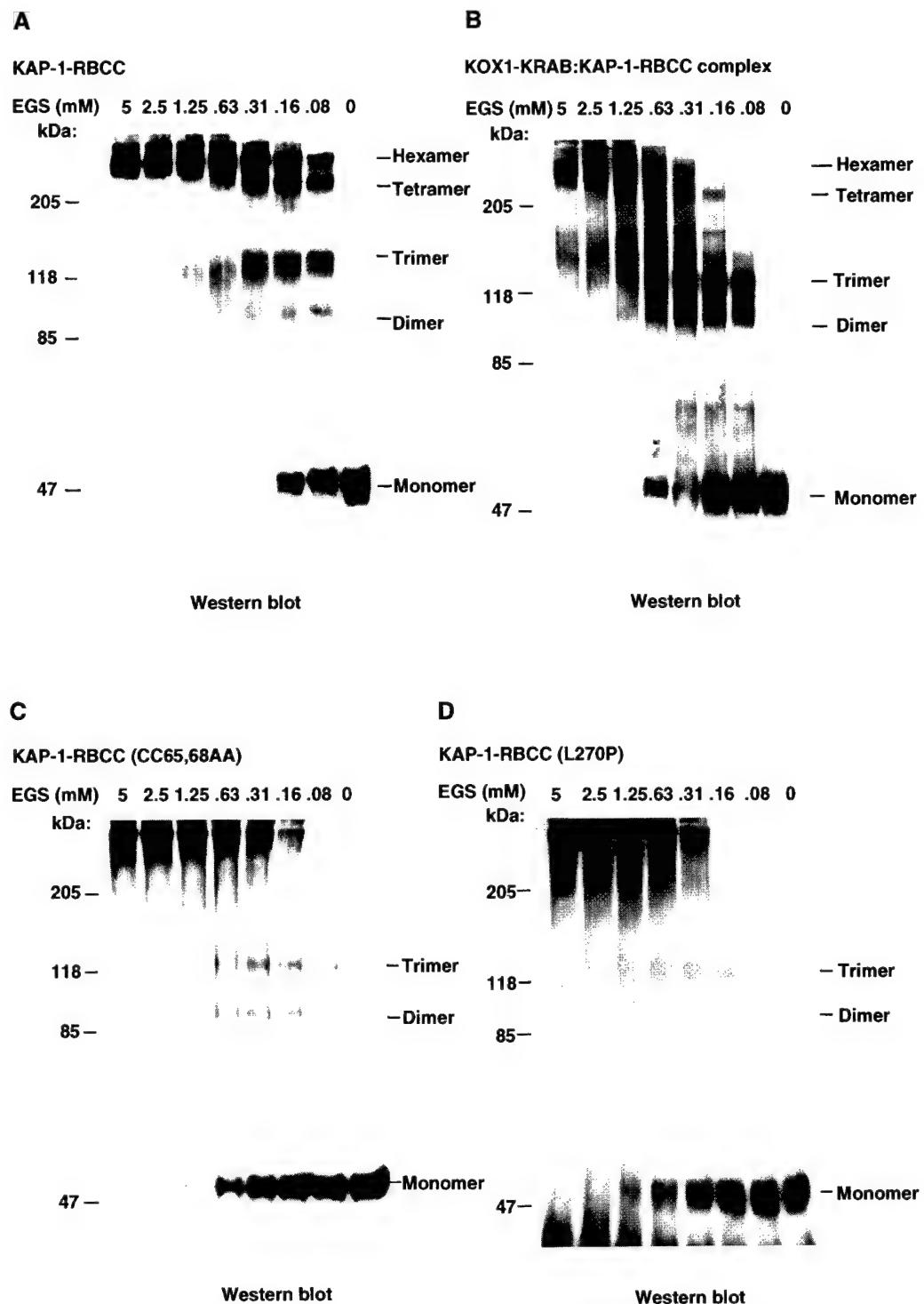


Figure 6. Chemical cross-linking analysis of the KAP-1-RBCC domain. Equal amounts ($5 \mu\text{M}$) of (a) wild-type KAP-1-RBCC, (b) a KOX1-KRAB:KAP-1-RBCC complex, (c) the (CC65,68AA) mutant, and (d) the (L270P) mutant were incubated with various concentrations of EGS. The cross-linked KAP-1-RBCC products were resolved by SDS-PAGE and visualized by Western blot analysis. An anti-His monoclonal antibody (Qiagen) was used to detect the KAP-1-RBCC proteins in panels (a), (c) and (d). An anti-KAP-1-RBCC rabbit IgG was used to detect the KAP-1-RBCC protein in (b). The position of molecular markers is indicated on the left of each Western blot. The estimated multi-meric forms of the proteins are indicated on the right of each Western blot.

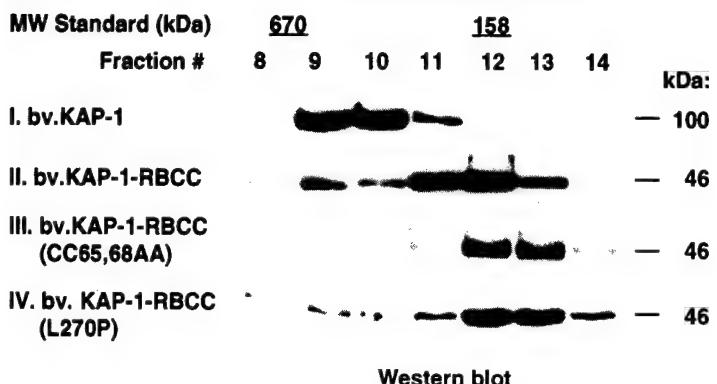
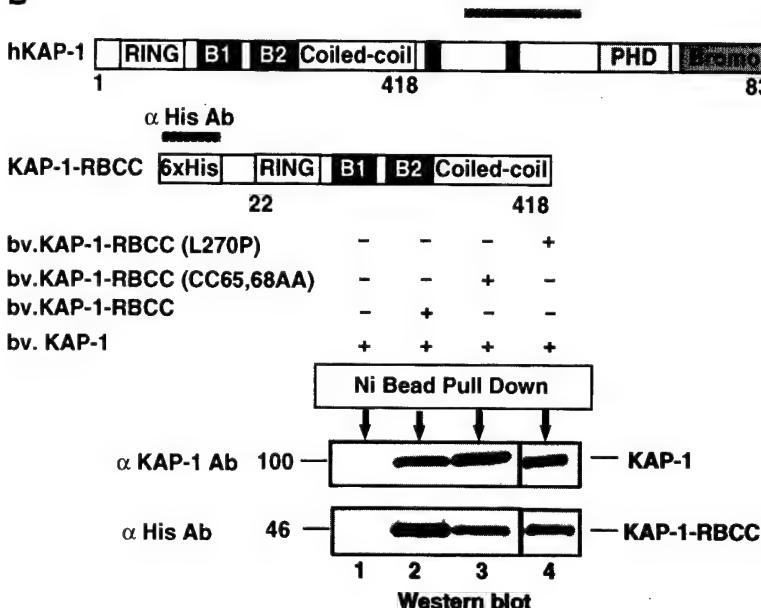
A GEL FILTRATION

B


Figure 7. Expression and functional analysis of KAP-1 proteins expressed from baculovirus vectors. (a) Gel-filtration analysis of full-length KAP-1 (I), wild-type KAP-1-RBCC (II), and mutant KAP-1-RBCC (III, IV). The elution of each protein from a Superdex 200 column was detected by Western blot analysis. The fraction numbers are labeled at the top of each lane, and the protein standards are indicated above the blots. An affinity-purified rabbit polyclonal anti-KAP-1 serum was used to detect full-length KAP-1. An anti-His monoclonal antibody was used to detect the KAP-1-RBCC proteins. (b) Formation and detection of KAP-1:KAP-1-RBCC hetero-oligomers *in vivo*. Diagrams of full-length KAP-1 and 6His KAP-1-RBCC proteins expressed in baculovirus are presented. The proteins were either expressed individually, or co-expressed in Sf9 cells. Hetero-oligomerization of the KAP-1-RBCC domain with full-length KAP-1 was detected by binding the complex to Ni-NTA resin. The bound proteins were detected by Western blot analysis. The bound full-length KAP-1 protein was detected using a affinity purified rabbit polyclonal anti-KAP-1 serum, and the KAP-1-RBCC proteins were detected with the anti-His monoclonal antibody. The -+ represents the viruses used to infect Sf9 cells.

yielded a supershift that was easily distinguishable from the supershift produced by bv.KAP-1-RBCC (Figure 8(a), lane 5). It is well established that this difference can be exploited as an assay for hetero-oligomerization of proteins by searching for intermediate mobility supershifts (Hope & Struhl, 1987). The bv.KAP-1 and bv.KAP-1-RBCC co-infected extracts produced a spectrum of five intermediate supershifts (Figure 8(a), lane 6), including the bands that co-migrate with the authentic bv.KAP-1 and bv.KAP-1-RBCC supershifts. This suggests that discrete hetero-oligomers composed of various combinations of full-length KAP-1 and wild-type RBCC domain can bind to the KRAB domain. Each of these supershifts contains either KAP-1, KAP-1-RBCC, or a combination thereof as determined by antibody supershift experiments (data not shown). If the complex between bv.KAP-1 and bv.KAP-1-RBCC was a simple hetero-dimer, a single intermediate species would be predicted. The presence of multiple discrete species strongly

suggests that hetero-oligomers are formed, most likely hetero-trimers. Co-expression of KAP-1 with either mutant RBCC protein yielded mobility shifts that were consistent only with the supershift observed for full-length KAP-1 (Figure 8(a), lanes 7 and 8). A plausible interpretation of these results is illustrated by Figure 8(b). Bands A and B represent hexamer and trimer forms of full-length KAP-1, respectively, which are in equilibrium; the hexamer is formed most efficiently by the full-length KAP-1 protein and binds with highest affinity to the KRAB domain. Bands C and D are hetero-trimers that contain one and two molecules of KAP-1-RBCC per complex, respectively. These complexes can bind the KRAB domain but may lack the ability to form a hexamer under these conditions. Band E is the homo-trimer of wild-type KAP-1-RBCC. The mutant RBCC proteins retain the ability to hetero-trimerize, but any complex containing a single mutant molecule would be inactive for KRAB domain binding, as evidenced by the loss of

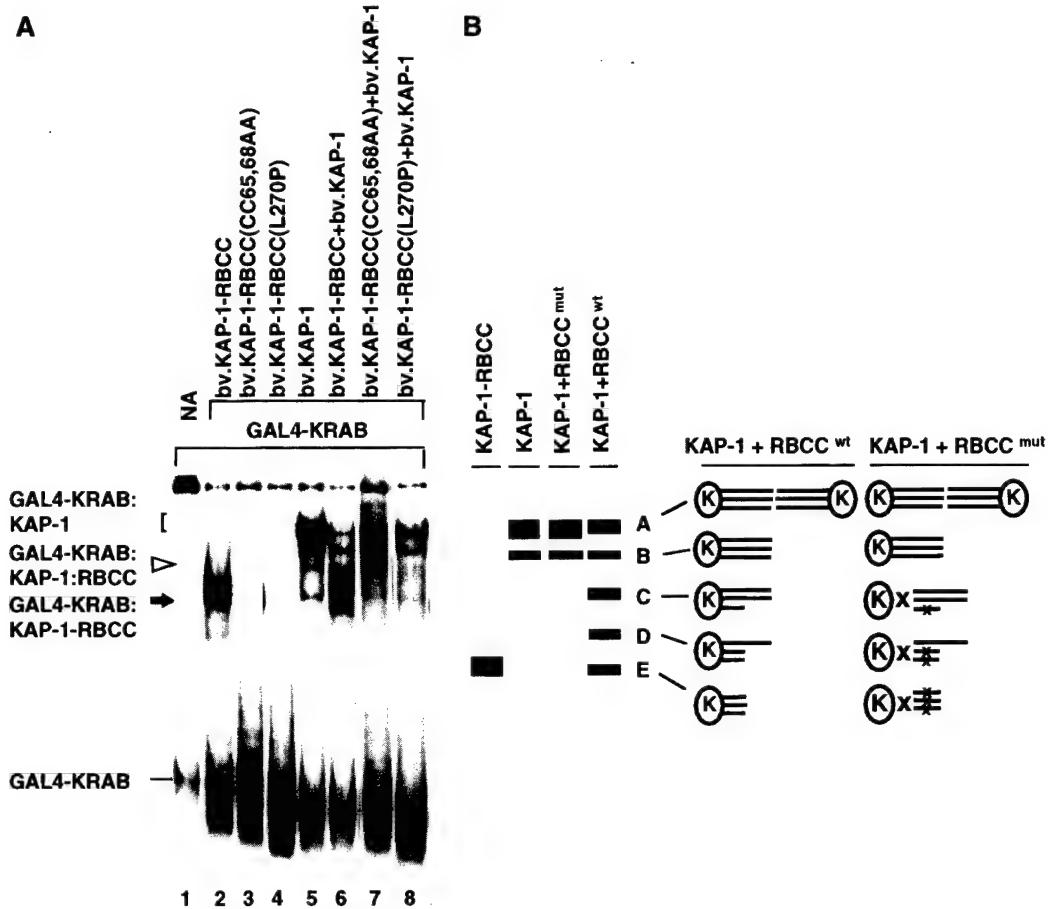


Figure 8. Binding of KAP-1-RBCC:KAP-1 hetero-oligomers to a DNA-bound KRAB domain *in vitro*. (a) The binding of KAP-1-RBCC:KAP-1 hetero-oligomers to GAL4-KRAB, as detected by EMSA. Each reaction contained a constant amount of purified GAL4-KRAB (100 ng), 5 µg of baculovirus whole-cell lysates, in which full-length KAP-1 protein was co-expressed with either the wild-type or mutant KAP-1-RBCC protein, and 1 µl of ³²P-labeled GAL4 probe. The symbols on the left side of the panel define the following complexes: dash, DNA:GAL4-KRAB alone; filled arrow, DNA:GAL4-KRAB:KAP-1-RBCC complex; open arrowhead, DNA:GAL4-KRAB:KAP-1-RBCC complex; bracket, DNA:GAL4-KRAB:KAP-1 complex. The free probe was allowed to run off the bottom of the gel for better resolution of the mobility shifts at the top of the gel. Note that no binding of the (CC65,68AA), or the (L270P) mutant KAP-1-RBCC proteins to GAL4-KRAB was detected (lanes 3 and 4). Also, hetero-oligomers of full-length KAP-1 with either of the mutant KAP-1-RBCC proteins failed to yield any intermediate supershifts (lanes 7 and 8) of the GAL4-KRAB complex, which were observed for hetero-oligomers of full-length with the wild-type KAP-1-RBCC domain (lane 6). (b) An interpretation for the formation of GAL4-KRAB:KAP-1:KAP-1-RBCC hetero-oligomers shown in the EMSA assay (a).

complexes C, D, and E in Figure 8(a), lanes 7 and 8. These results suggest that oligomerization alone is not sufficient for interaction of KAP-1 with the KRAB domain. Moreover, individual components within the RING finger, B boxes, and coiled-coil region independently assist the proper oligomerization of KAP-1 and its interaction with the KRAB domain.

Heterologous oligomerization domains fail to reconstitute the KOX1-KRAB:KAP-1 interaction

The data presented above suggest a simple model for the KAP-1-RBCC domain: that the coiled-coil region is a parallel trimeric association that serves to properly orient the RING finger and

B boxes for binding to the KRAB domain. If this hypothesis is correct, we may be able to substitute a similarly structured coiled-coil domain to reconstitute KRAB binding. We thus substituted the coiled-coil region of KAP-1 with well-characterized heterologous oligomerization domains for trimerization (HSF) (Drees *et al.*, 1997), dimerization (GCN4-LZ), and tetramerization (GCN4-TZ) (Waterman *et al.*, 1996) (Figure 1(d)). These heterologous KAP-1-RBCC fusion proteins were expressed as soluble proteins and were properly oligomerized, as determined by chemical cross-linking and gel filtration (data not shown). As illustrated in Figures 1(d) and 9, each of these proteins, KAP-RB-[HSF_{Tri}], KAP-RB-[GCN4_{Dim}], and KAP-RB-[GCN4_{Tet}], failed to interact with the

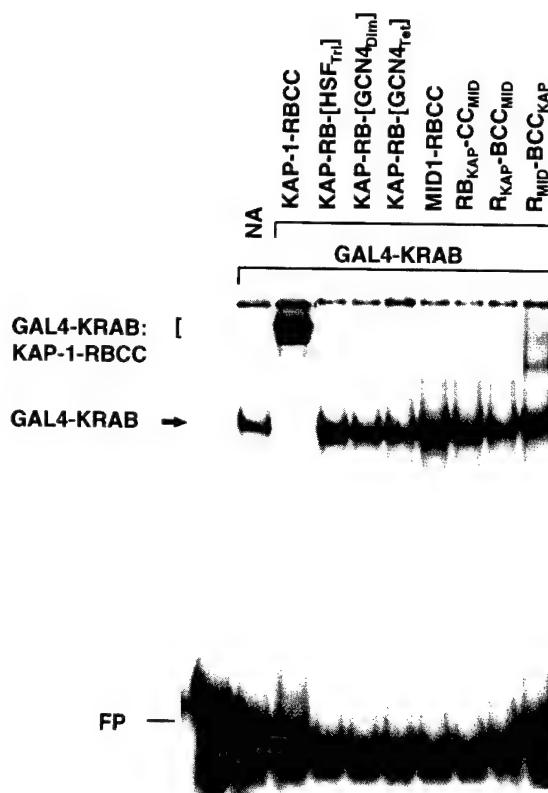


Figure 9. Substitution of the RING finger, B-boxes, and the coiled-coil sub-domains of KAP-1 disrupts the KOX1-KRAB:KAP-1 interaction: 1 µg of each wild-type KAP-1-RBCC or heterologous KAP-1-RBCC fusion protein was tested for its ability to interact with the KRAB domain *in vitro*, via the EMSA. The wild-type KAP-1-RBCC protein efficiently supershifted the DNA-bound GAL4-KRAB protein to form a complex indicated by the bracket. None of the heterologous RBCC fusion proteins demonstrated this activity, although the R_{MID}-BCC_{KAP} fusion protein showed a little of this activity. The filled arrow indicates the DNA:GAL4-KRAB mobility shift. FP, free probe.

DNA-bound GAL4-KRAB protein. Thus, the KAP-1 coiled-coil region may provide a unique oligomerization surface required for KRAB binding or contribute directly to KOX1-KRAB domain recognition *via* specific contacts.

The studies described here suggest that the RING finger, B boxes, and coiled-coil region collectively appear to constitute a tripartite interface for protein-protein interactions that is conserved in a number of cellular proteins including TIF1 α , TIF1 γ , and MID1 (Figure 1). To more precisely address the specificity of each motif in this apparent protein-protein interaction domain of KAP-1, we have substituted the RING finger, B boxes, and coiled-coil sub-domains of KAP-1 with the corresponding motif from MID1 (Quaderi *et al.*, 1997) (Figure 1(d)). We expressed and purified the MID1-RBCC domain and heterologous KAP-1/

MID1-RBCC fusion proteins from bacteria (data not shown), and tested their ability to interact with the KRAB domain. In general, the MID1-RBCC domain and each of the KAP-1/MID RBCC fusion proteins failed to bind to the DNA-bound GAL4-KRAB protein. However, R_{MID}-BCC_{KAP} fusion protein showed weak potential to shift the DNA:GAL4-KRAB binary complex (Figures 1(d) and 9).

Together, these data suggest that the RBCC domain of KAP-1 functions as an integrated structural unit to independently facilitate the oligomerization of KAP-1 and its highly specific interaction with the KRAB domain.

Discussion

The RBCC tripartite domain has emerged as a highly conserved module in proteins that play causal roles in cellular differentiation and development, and whose alteration can result in human disease (El-Husseini & Vincent, 1999; Ogawa *et al.*, 1998; Perrin & Lacroix, 1998; Tatematsu *et al.*, 1998). The spatial conservation of sub-domains in the RBCC motif and their presence in large cellular complexes has led to the suggestion that the RBCC domain mediates specific protein-protein interactions. However, these targets, even for the most well-characterized RBCC proteins have eluded investigators.

Here, we present the first demonstration that a highly specific, physiologically relevant protein-protein interaction can be reconstituted with a purified RBCC domain and its corresponding binding partner, thus providing proof that the RBCC domain is a protein-protein binding surface. We can draw the following conclusions from our data: (1) the interaction is direct; it does not apparently require post-translational modification and is stable to *in vitro* biochemical manipulation. (2) A stable ternary complex can be formed between DNA, a DNA-bound KRAB domain and the KAP-1-RBCC protein; thus, the KRAB-KAP-1 interaction does not inhibit DNA-binding in the system we have used. (3) Recognition of the KRAB domain by the KAP-1-RBCC is highly specific, as the related RBCC proteins MID1 and TIF1 α do not bind under these experimental conditions. (4) Binding requires structural elements of all the sub-domains of the KAP-1-RBCC, as mutation of each individual sub-domain has the potential to abolish KRAB domain binding. (5) Each RBCC sub-domain may cooperate in providing specificity for interaction, as individual sub-domain swaps with related RBCC proteins fail to recapitulate binding. (6) The KAP-1-RBCC domain must properly homo-oligomerize in order to bind the KRAB domain: this oligomerization state is probably a trimer and further association to a hexamer may occur. (7) Oligomerization is necessary but not sufficient for KRAB domain binding, implying that separate functional domains for oligomerization and KRAB recognition exist.

(8) The coiled-coil sub-domain itself provides critical binding and/or determines orientational specificity for KRAB domain binding, as heterologous oligomerization domains fail to recapitulate binding. The picture that emerges is that the KAP-1-RBCC domain forms an integrated unit composed of essential sub-domains that probably function cooperatively to recognize a protein ligand with high specificity.

Although it has been demonstrated that many RBCC proteins oligomerize, its role in ligand binding has not been determined. Our data with the KRAB-KAP-1 system firmly establish that oligomerization is required for binding. A major issue is the oligomerization of KAP-1, before and after KRAB binding and whether binding stabilizes the oligomeric species. We favor the interpretation that the KRAB-bound KAP-1-RBCC domain is a trimer or hexamer based upon a number of observations. Analysis of the predicted coiled-coil region using the multi-coils program predicts the potential for two independent coiled regions with the highest probability that each forms a trimer. Gel filtration of soluble *E. coli* KAP-1-RBCC in the presence or absence of the KRAB domain shows a complex of apparent molecular mass three times greater than the monomeric molecular mass. Chemical cross-linking in solution shows apparent trimeric and hexameric forms of the KAP-1-RBCC as the predominant species, and the production of the trimeric cross-linkable form is greatly enhanced by pre-incubation with recombinant KRAB domain. Analytical ultracentrifugation experiments show mixtures of monomers, trimers, and hexamers. Together, these observations are consistent with the potential for a monomer-trimer-hexamer equilibrium. The KAP-1-RBCC protein also sedimented as an apparent monomer in standard velocity sedimentation in sucrose gradients, suggesting that oligomers were not stable under these conditions. However, pre-incubation of the KAP-1-RBCC with the KRAB domain resulted in a sucrose gradient sedimentation profile of the complex consistent with a trimer. The shift from apparent monomer to apparent trimer in sedimentation is unlikely to be accounted for by simple incorporation of multiple molecules of the KRAB domain into the complex, since preliminary stoichiometry estimates suggest that the KAP-1-RBCC:KRAB complex is approximately 3:1 (unpublished results). In summary, the behavior of the complex in these assays most likely indicates KRAB-induced trimer formation or a conformational change detected by sedimentation.

We confirmed these studies using KAP-1-RBCC and full-length KAP-1 protein expressed in eukaryotic cells. The co-infection experiments clearly show that hetero-oligomers can be efficiently formed between full-length KAP-1 and KAP-1-RBCC, and that the hetero-oligomers are functional for KRAB binding. Most importantly, three intermediate supershifts were observed. This strongly suggests that the active binding species is oligomeric, most likely a trimer or hexamer.

The mutations in either the RING finger (CC65,68AA) or the coiled-coil region (L270P) when expressed individually in Sf9 cells were inactive in KRAB domain binding. Thus, either they are defective in homo-oligomerization or form an oligomer unproductive for KRAB binding. Both of these mutant proteins show slightly smaller apparent molecular masses in gel filtration, suggesting an unfolded state. This apparent monomeric species cannot be converted to the apparent oligomeric species (as detected by sucrose gradient sedimentation or cross-linking) via incubation with the KRAB domain. Thus, the apparent KRAB-induced oligomerization is required for binding and both the RING finger and the first coiled-coil are important for this oligomerization.

However, it is puzzling that these mutants, when co-expressed with full-length KAP-1 in baculovirus-infected cells efficiently form a hetero-oligomeric complex. Clearly, these hetero-oligomeric complexes are not competent for KRAB domain binding, since there were no intermediate supershift bands observed from a mutant/wild-type co-infected extract. Thus, either the binding is non-specific, or the full-length KAP-1 protein compensates for an oligomerization defect in the mutant RBCC protein.

Hetero-oligomers that contain one mutant sub-unit are apparently defective in KRAB binding and thus these hetero-oligomers appear to act as true dominant-negatives. This behavior has been observed in many other oligomeric proteins, where functionally-separate oligomerization and ligand binding domains are present (Hurst, 1995; Baxevanis & Vinson, 1993). An important point is that the RING finger mutant (CC65,68AA) fails to form a productive homo-oligomer or hetero-oligomer with wild-type KAP-1, even though it contains an intact coiled-coil region. This observation clearly demonstrates that the RING finger participates both in oligomerization and ligand recognition.

Our mutational analysis of the individual RBCC sub-domains clearly shows that the core structural integrity of the RING finger appears to be required for KRAB domain binding. Many engineered and naturally occurring RING finger mutations have been studied in other proteins but none has yet been correlated with alteration of a specific RING-protein interaction in a purified system. Homo or hetero-oligomerization via RING-RING interaction has been observed in PML, BRCA1, Bmi-1, Mph2 (Borden *et al.*, 1995a; Hemenway *et al.*, 1998; Meza *et al.*, 1999). The RING finger mutant BRCA1 proteins show altered oligomerization, subcellular localization, interaction with protein partners, and ability to suppress cell growth (reviewed by Irminger-Finger *et al.*, 1999). RING finger mutations in the proto-oncogene PML have profound effects on the localization of the protein to the ND10 nuclear bodies (Borden *et al.*, 1995a), which have been correlated with effects on both cell survival mediated by wild-type PML (Borden *et al.*, 1997) and oncogenesis as mediated by the

PML-RAR α fusion (Fagioli *et al.*, 1998). Finally, the RING finger of RFP is critical to its cell transformation function as the rfp-ret fusion oncogene (Cao *et al.*, 1997, 1998; Hasegawa *et al.*, 1996).

The mutational analysis of the B1 and B2 boxes of KAP-1 has also been informative. Interestingly, the triple-cysteine substitution in the NH₂-terminal B1 box had no effect on KRAB domain binding, whereas a similar substitution in the analogous cysteine residues in the B2 box completely abolished KRAB domain binding. If these substitutions abolished the structural integrity of both B boxes, it is reasonable to conclude that there is an unequal contribution of each B-box to KRAB domain binding and/or KAP-1 oligomerization. The evidence that the B-box is required for the proper orientation of the coiled-coil region (Cao *et al.*, 1997) could explain why mutation of the B2 box adjacent to the coiled-coil in KAP-1 abolished KRAB binding, whereas mutation of the distal B1 box had no effect.

Oligomerization mediated by the coiled-coil region is one of the most important determinants of the biological properties of RBCC proteins, and is likely responsible for their legendary potential for aggregation and multimerization *in vitro*. The L270P substitution in the KAP-1-RBCC expressed in *E. coli* showed aberrant sedimentation and failure to oligomerize in the chemical cross-linking assay (in the absence of the KRAB domain). These properties correlated with an impaired ability to bind GAL4-KRAB, although significant binding was detected at very high levels of input KAP-1-RBCC (L270P). However, baculovirus-expressed KAP-1-RBCC (L270P) was completely inactive for GAL4-KRAB binding, even at very high concentrations. Surprisingly, the bv.KAP-1-RBCC (L270P) still efficiently bound to wild-type full-length bv.KAP-1 in co-infected cells. However, these hetero-oligomeric complexes were not able to bind the KRAB domain, thus exhibiting an apparent dominant-negative activity essentially identical with the KAP-1-RBCC, RING finger mutant (CC65,68AA) discussed above. Like the RING finger mutant, we see a difference in homo-oligomeric *versus* hetero-oligomeric potential with wild-type KAP-1 in the KAP-1-RBCC (L270P) mutant. The L306P substitution in coil II is clearly different. The *E. coli*-produced L306P protein is completely blocked in its ability to bind the KRAB domain, even at high concentration. Thus, either the L270 and L306 leucine residues that we targeted are not equivalent in terms of their contribution to each coiled-coil structure, or coil I and coil II are not equivalent in their contribution to either oligomerization or KRAB domain recognition. It is a distinct possibility that one (or both) coils participate directly in KRAB domain recognition. In summary, our mutational analysis for the KAP-1-RBCC sub-domains strongly suggests that each sub-domain contributes to KRAB domain binding, either as a

direct recognition element or indirectly by mediating appropriate homo-oligomerization.

To begin distinguishing these possibilities, we constructed a number of fusion proteins containing heterologous oligomerization domains and "swapped" RBCC sub-domains from KAP-1 orthologues. Our strategy of replacing the coiled-coil region of KAP-1 with heterologous oligomerization domains was based on the following plausible model for the KAP-1-RBCC oligomer: that the coiled-coil region forms a parallel trimeric association of helices that orient the RING finger and B-boxes for KRAB binding. If the coiled-coil serves only as an oligomerization motif, it would be predicted that it could be substituted for by a similar motif, as has been shown in other oligomeric proteins. The best evidence that supports such a potential model comes from recent structural analyses of the TRAF proteins (McWhirter *et al.*, 1999; Park *et al.*, 1999), which encode RING finger and coiled-coil regions and are a subfamily of the RBCC proteins. The TRAFs homo and hetero-trimerize *via* a parallel association of the coiled-coil region. Trimerization properly orients the TRAF domain for binding to the cytoplasmic receptor, and this trimerization function can be substituted by heterologous trimerization domains (McWhirter *et al.*, 1999; Park *et al.*, 1999). However, KAP-1-RB proteins fused to heterologous dimerization, trimerization, or tetramerization domains completely failed to bind the KRAB domain. Our alternative approach, using domain swaps from the MID1-RBCC protein yielded similar, uniformly negative results. That is, substitution of the RING finger, B-box-coiled-coil, or coiled-coil region alone from MID1 for the corresponding region in KAP-1 failed to recapitulate KRAB domain binding. Together, these results with chimeric RBCC proteins strongly suggest that the RBCC domain of KAP-1 binds as an integrated functional unit to its ligand, and that each sub-domain contributes to protein recognition and/or oligomerization. It is noteworthy that the intact RBCC domains from three independent RBCC proteins, TIF1 α , MID1 and TIF1 γ (Venturini *et al.*, 1999) all fail to bind the KRAB domain. These observations underscore the apparent high degree of specificity for ligand recognition that is built into the RBCC domain of KAP-1 and probably other RBCC proteins.

An important point is whether endogenous RBCC proteins form hetero-oligomers *via* RBCC-RBCC interactions. These interactions would serve to increase the combinatorial complexity for target recognition by hetero-oligomeric RBCC complexes. It is interesting to note that apparent hetero-trimerization has been observed among certain members of the TRAF subfamily of RBCC proteins, although it is unclear whether this alters the biological signaling properties/specificities of these molecules (Pullen *et al.*, 1999). Even among the clearly paralogous TIF1 family of RBCC proteins, only indirect weak binding between family members has been detected (Venturini *et al.*, 1999). Rather remarkably,

the MID1 and MID2 RBCC proteins, which show overall 77% amino acid sequence identity, fail to hetero-dimerize *in vivo* (Buchner *et al.*, 1999; Cainarca *et al.*, 1999). However, there is indirect evidence that RFP can hetero-oligomerize and co-localize with PML, and that this activity requires the B box and coiled-coil regions of RFP (Cao *et al.*, 1998). Clearly, studies with purified components as demonstrated here, are required to resolve this important issue.

Our primary goal in this study was to begin to establish a set of principles for oligomerization and ligand recognition for the RBCC family of proteins using the KRAB-KAP-1 system as a model. However, equally important are the implications that our findings have for defining the biological functions of KAP-1 as a corepressor *in vivo*. The KRAB domain is currently represented in more than 100 proteins, most of which are C₂H₂ zinc-finger proteins and are likely to function as gene-specific repressors. Although we have exclusively utilized the KRAB domain of KOX1, more than ten independent KRAB domain proteins have been shown to bind KAP-1 and thus it is a reasonable assumption that the vast majority of this family will utilize KAP-1 as their corepressor. Our current working model for KRAB-ZFP-KAP-1-mediated repression is illustrated in Figure 10. A KRAB-ZFP binds DNA sequence-specifically through its array of C₂H₂ zinc fingers. The DNA-bound KRAB domain recruits the KAP-1 corepressor *via* a direct interaction with the RBCC domain. Binding may induce oligomerization of KAP-1, which is required, or KAP-1 may be pre-existing in an oligomeric state. Binding of KAP-1 by the KRAB-ZFP in proximity to the transcription start site is necessary but not sufficient for repression. The KAP-1 repression domains comprised of the HP1 binding domain, the PHD finger and the bromodomain, which

together function to assemble the machinery required for repression. A major goal has been to define this machinery "downstream" of KAP-1, which is required and, in light of the evidence presented here, whether KAP-1 oligomerization also plays a role in this recruitment. A likely direct target of KAP-1 in this context was recently defined by us and others as the heterochromatin protein HP1 (Ryan *et al.*, 1999). The HP1 proteins are a small family of low molecular mass, non-histone chromosomal proteins, some of which are tightly associated with silenced heterochromatin. We have shown that KAP-1 directly binds to purified recombinant HP1 protein, and that a stable quaternary complex can be formed between DNA, a KRAB-ZFP, KAP-1 and HP1. Mutations in KAP-1 that abolish HP1 binding significantly impair KAP-1-mediated repression. Furthermore, HP1 and KAP-1 extensively colocalize to heterochromatic regions in interphase nuclei. Thus, it is tempting to speculate that KAP-1 serves to nucleate HP1-mediated stable repression on a template, and that the oligomerization of KAP-1 as it binds the KRAB domain has a role in the nucleation of heterochromatin.

In summary, we have provided the first analysis of a direct protein-protein interaction involving an RBCC protein and its physiologically relevant protein ligand. These studies have highlighted the apparently high specificity for recognition inherent in this protein-protein interface. Further refinement of these macromolecular interactions could lead to strategies for harnessing KAP-1-mediated repression *in vivo* using synthetic transcription factors for possible therapeutic benefit. One could also envision a small-molecule inhibitor strategy directed at the KRAB-KAP-1 system, which may be useful for reprogramming the expression of stably repressed genes and thus altering cellular phenotypes.

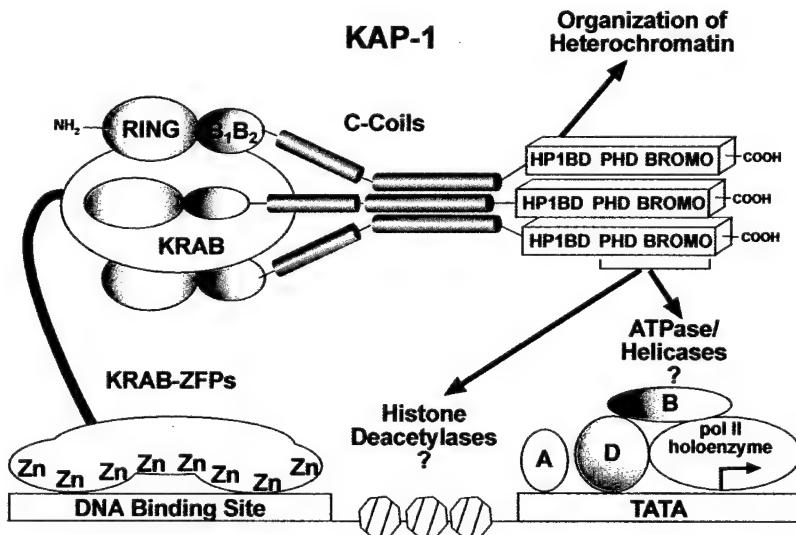


Figure 10. Model for the KRAB-KAP-1 interaction. This model is described in Discussion.

Materials and Methods

Preparation of plasmids

The plasmid expressing the 6His GAL4-KRAB (1-90) protein was constructed *via* polymerase chain reaction (PCR) using the pM1-KOX1, 1-90 plasmid (Margolin *et al.*, 1994) as a template. A 5' oligonucleotide, which incorporated a *Bam*HII site immediately 5' to the GAL4 initiator methionine residue and a 3' oligonucleotide, which incorporated a stop codon and a *Hind*III site after amino acid residue 90 of KOX1, were used to amplify the desired sequence. The digested PCR product was cloned into the corresponding restriction sites of the pQE30 vector (Qiagen).

The plasmid expressing the 6His KRAB (1-90) protein was produced by subcloning the DNA segment encoding residues 1-90 of KOX1-KRAB from pM1 into the pSP73 vector, using the 5' *Eco*RI and 3' *Xba*I sites. This fragment was then excised and subcloned *via* the flanking *Bam*HII and *Pst*I sites in pSP73 into the corresponding sites of the pQE30 vector. The protein expressed from this plasmid contained the 18 NH₂-terminal amino acid residues, MRGSHHHHHGSDIIDEF, followed by residues 1-90 of the KOX1 protein, then the C-terminal sequence RVDLQPSLIS is present.

The pQE30 KAP-1-RBCC plasmid was constructed by amplifying the DNA fragment encoding residues 22-418 of human KAP-1 using the primer pair, forward 5' CGT TGA ATT CTA CCA TGG CTT CGT GGG GAT CCC 3', reverse 5' CGT AAA GCT TGC CTG TTG AGT TAG TGC C 3'. The PCR product was cleaved with *Eco*RI and *Hind*III, and cloned into pBluescript II KS +/− to generate pBS-KS KAP-1-RBCC. This fragment was then excised with the flanking 5' *Sma*I site and 3' *Hind*III sites in pBluescript II KS +/− and cloned into the pQE30 plasmid. The expressed KAP-1-RBCC protein contained the 18 NH₂-terminal amino acid residues, MRGSHHHHHGSAELGT, followed by residues 22-418 of the KAP-1 protein, then the C-terminal sequence KLN. The point mutations within the RBCC domain of KAP-1 were generated in this plasmid using PCR-mediated mutagenesis. The mutagenic primers contained the following codons: (CC65,68AA), TGC to GCC; (CCC152,153,156AAA), TGC to GCC, TGT to GCT; (C209A,H212A), TGC to GCC, CAC to GCC; (L270P), TTG to CCG; (L306P), CTG to CCG; (V260N), GTG to AAC; (V260D), GTG to GAC.

A baculovirus vector (pVL1392) encoding residues 22-418 of KAP-1 was generated by first digesting pBS-KS KAP-1-RBCC with *Hind*III. The overhang in the linearized plasmid was filled in with Klenow enzyme, and then the DNA was digested with *Eco*RI. This fragment was cloned into pVL1392 (PharMingen) at the *Eco*RI and *Sma*I sites. The protein expressed contained the 16 NH₂-terminal amino acid residues, MASWGSHHHHHHDPGY, followed by residues 22-418 of the KAP-1 protein, then the C-terminal sequence KLGSPPGTRQEPKTHSLQGNP.

The pQE30 TIF1α-RBCC plasmid was produced by amplifying the DNA fragment encoding residues 1-414 of TIF1α using the primer pair, forward 5' GCA TAG ATC TAT GGA GGT GGC TGT GGA GAA G 3'; reverse 5' CGA TGC ATG CTC AGA GCC CAG AAA CTA GGA TC 3'. The PCR product was digested with *Bgl*II and *Sph*I, and cloned into the *Bam*HII and *Sph*I sites of the pQE30 vector.

The plasmids expressing the KAP-1 RB region fused to various heterologous oligomerization domains and

KAP-1/MID-1 domain swaps were generated by overlap-extension PCR-mediated mutagenesis. The GCN4-LZ and GCN4-TZ domains have been described and spanned residues 253-281 and 250-281 of GCN4, respectively (Waterman *et al.*, 1996). The *Saccharomyces cerevisiae* heat-shock transcription factor (HSF) trimerization domain spanning residues 321-441 was used (Drees *et al.*, 1997). The forward primers are: KAP-RB-[HSF_{Trl}], KAP-RB-[GCN4_{Dim}], KAP-RB-[GCN4_{Tet}], RB_{KAP}-CC_{MID}, R_{KAP}-BCC_{MID}, 5' AGC GGA TCC GAA TTC ACC ACC ATG GCT CAC CAT CAC CAT CAC 3' (*Bam*HII); MID-RBCC, R_{MID}-BCC_{KAP}, 5' GCA TAG ATC TAT GGA AAC ACT GGA GTC AG 3' (*Bgl*II). The reverse primers are: KAP-RB-[HSF_{Trl}], 5' CGT TAA GCT TTC AAG TGC CTG TGT TGT CGC G 3' (*Hind*III); KAP-RB-[GCN4_{Dim}], 5' CGT TAA GCT TTC AGC GTT CGC CAA CTA A 3' (*Hind*III); KAP-RB-[GCN4_{Tet}], 5' CGT TAA GCT TTC AAC GTT CAC CCA ATA ATT T 3' (*Hind*III); MID-RBCC, 5' CGA TGC ATG CTC AAT CTA AGG CAA AGG TGT C 3' (*Sph*I); RB_{KAP}-CC_{MID}, R_{KAP}-BCC_{MID}, 5' GAT CAA GCT TTC AAT CTA AGG CAA AGG TGT C 3' (*Hind*III); R_{MID}-BCC_{KAP}, 5' ATG CAA GCT TTC ATG TTG AGT TAG TGC CAG GAC 3' (*Hind*III). Fusion primers (forward) are: KAP-RB-[HSF_{Trl}], 5' TGA GGA ACC AGC GCA AGA ATG GTC ATT TAT TGC AGG 3'; KAP-RB-[GCN4_{Dim}], 5' TGA GGA ACC AGC GCA AGA TCC GCG GGC GTC ACC TT 3'; KAP-RB-[GCN4_{Tet}], 5' TGA GGA ACC AGC GCA AGA TCC GCG GTA ATC GTC TG 3'; RB_{KAP}-CC_{MID}, 5' GTG AGG AAC CAG CGC AAG AAG CAA AAC TTA GAG AGT AAC 3'; R_{KAP}-BCC_{MID}, 5' AAT TAT TTC ATG CGT GAT AGT GGG CTC AAG CGC AAC GTC ACC 3'; R_{MID}-BCC_{KAP}, 5' CTC AGC CAG CGA GGT CTA GAC GGC AGC AAG GCT GCC ACC GAC 3'. The second-round PCR products were digested with the appropriate enzymes defined for each fusion in parentheses above and subcloned into compatible restriction sites of the pQE30 vector. All plasmids generated by PCR were confirmed by DNA sequencing on both strands to verify the presence of desired mutations and appropriate fusions (Wistar DNA Core Facility).

Protein purification

Escherichia coli SG13009 cells (Qiagen) bearing the desired plasmid were propagated with aeration at 37°C in one liter of Luria broth to an *A*₆₀₀ of approximately 0.6. IPTG was added to 1 mM, and growth at 37°C was continued for four hours. The cells were harvested by centrifugation.

The purification of KOX1-KRAB protein was performed at room temperature (Shi *et al.*, 1997). For one liter of cell culture, the cell pellet was resuspended in 40 ml of P300 buffer (10 mM Na₂HPO₄, 1.4 mM KH₂PO₄, 2.7 mM KCl, 450 mM NaCl, pH 7.0, 1 mM phenylmethylsulfonyl fluoride (PMSF)). The cells were lysed by sonication on ice and the cell extract was centrifuged at 12,000 g for 20 minutes. The pellet was dissolved in 40 ml of lysis buffer (20 mM Tris-HCl (pH 7.8), 0.5 M NaCl, 5 mM imidazole, 8 M urea), and incubated for one hour. The lysate was then centrifuged at 12,000 g for 20 minutes at 4°C. To the supernatant, 1-1.5 ml packed column volume of Ni-NTA resin (Qiagen), previously equilibrated in lysis buffer, was added in batch and the incubation continued with rotation for one hour. The resin was then washed with lysis buffer and loaded onto a disposable column, followed by an additional volume of wash buffer (20 mM Tris-HCl (pH 7.8), 0.5 M NaCl, 20 mM imidazole, 8 M urea). The KRAB protein bound

to the nickel column was renatured on the column by stepping the urea in lysis buffer from 8 M to 0 M. The renatured KRAB protein was eluted with elution buffer (20 mM Tris-HCl (pH 7.8), 0.5 M NaCl, 0.3 M imidazole). The protein was dialyzed against 20 mM Tris-HCl (pH 8.0), 0.5 M NaCl, followed by a second dialysis reducing the salt to 50 mM NaCl.

The KAP-1-RBCC proteins, mutants thereof, and the chimeras, were purified using both denaturing and non-denaturing conditions. The purification was initially performed at 4°C under non-denaturing condition as recommended by the manufacturer (Qiagen). The bacterial pellet was resuspended in P300 buffer (pH 7.0), 10% (v/v) glycerol, 10 mM imidazole, and lysed by sonication. The cell extract was centrifuged at 12,000 g for 20 minutes. The insoluble pellet was then used for purification under denaturing conditions (see below). The supernatant fraction containing soluble protein was incubated in batch with Ni-NTA resin for one hour. The resin was washed with P300 buffer. Then the resin was loaded into a disposable column, and washed with P300 buffer at pH 6.0. The protein was eluted with P300 (pH 7.0), 10% glycerol, 300 mM imidazole. The eluted protein was dialyzed against three changes of P300 (pH 7.0), 10% glycerol, 20 μM ZnSO₄, 0.5 mM DTT.

The insoluble proteins from the pellet fraction were purified and eluted from the Ni-NTA with 300 mM imidazole and low pH under denaturing condition as recommended by the manufacturer. The eluted proteins were diluted to 20 ml with P300, and dialyzed against P300 (pH 7.0), 10% glycerol, 20 μM ZnSO₄, 0.5 mM DTT. After dialysis, the protein was clarified by centrifugation and concentrated in an Amicon concentrator (Amicon) following the manufacturer's instruction.

Electrophoretic mobility shift assay (EMSA)

EMSA assays were performed essentially as described (Ryan *et al.*, 1999; Fredericks *et al.*, 1995). The purified recombinant GAL4-KRAB or GAL4-KRAB (DV-AA) protein was incubated in binding buffer with either purified bacterial KAP-1-RBCC, KAP-1-RBCC mutant/fusion proteins, or baculovirus whole-cell lysates for 15 minutes at 30°C, and then 1 μl of ³²P-labeled GAL4 probe (10⁵ cpm/μl) was added and the reaction was incubated for an additional 15 minutes at 30°C. The DNA:protein complexes were resolved on native polyacrylamide gels by electrophoresis at 450 V for one hour and 45 minutes in 45 mM Tris-borate (pH 8.3), 1 mM EDTA buffer at 4°C. The EMSA gels were dried and visualized by autoradiography. The GAL4 was the double-stranded synthetic oligonucleotide 5' GAT CCC GGA GGA CAG TAC TCC GT 3', which was labeled with [³²P]ATP as described (Fredericks *et al.*, 1995).

Reconstitution of the KOX1-KRAB:KAP-1-RBCC complex *in vitro*

To reconstitute the KOX1-KRAB:KAP-1-RBCC complex *in vitro*, highly purified bacterial KOX1-KRAB and KAP-1-RBCC proteins were incubated at a 1:1 molar ratio in buffer containing 20 mM Hepes, 300 mM NaCl, 0.5 mM DTT, 10% glycerol at 30°C for 30 minutes. The samples were chilled at 4°C for five minutes, and then centrifuged at 10,000 g for five minutes at 4°C. The supernatant was analyzed by gel-filtration, sucrose gradient sedimentation, and chemical cross-linking.

Gel-filtration analysis

Gel-filtration of highly purified proteins was performed using a Superdex 200 HR 10/30 column (Pharmacia) equilibrated in P300 buffer (pH 7.0), 10% glycerol, 0.1% NP-40, 20 μM ZnSO₄, 1 mM DTT, using an FPLC system (Pharmacia). The column was run at 4°C at a flow-rate of 0.5 ml/minute and 1 ml fractions were collected. Aliquots of each fraction were analyzed by SDS-PAGE or Western blot (Ryan *et al.*, 1999).

Sucrose gradient sedimentation

Purified bacterial KAP-1-RBCC protein (60 μg), KOX1-KRAB (60 μg) or, KOX1-KRAB (60 μg):KAP-1-RBCC (200 μg) complex in 200 μl of sucrose gradient buffer (10 mM sodium phosphate (pH 7.0), 140 mM NaCl, 0.5% NP-40, 1 mM PMSF, 2 μg/ml aprotinin, 3 μg/ml leupeptin) was layered onto a discontinuous 5% to 60% (w/v) sucrose gradient. The gradients were pre-formed at room temperature for one hour, followed by an additional hour at 4°C prior to loading the sample. The sucrose gradients were centrifuged for 16 hours at 4°C in an SW-50.1 rotor at 190,000 g. Approximately 20 fractions (250 μl each) were collected from the base of each centrifuge tube. Aliquots of each fraction were subjected to SDS-PAGE, and the proteins were stained with Coomassie brilliant blue.

Non-denaturing (ND)-PAGE

ND-PAGE was conducted by the procedure described in the ND protein Molecular Weight Determination Kit from Sigma. Purified bacterial KAP-1-RBCC protein (10 μg) and 10 μg of each protein standard were analyzed by ND-PAGE on gels of various polyacrylamide concentrations (6, 7, 8, and 10%), and the proteins were stained with Coomassie brilliant blue.

Analytical ultracentrifugation

The KAP-1-RBCC (0.7 to 2.0 mg/ml) sample was prepared for analytical ultracentrifugation by dialysis against a P300 buffer (pH 7.0), 10% glycerol, 20 μM ZnSO₄, 1 mM DTT. The sedimentation equilibrium experiments were performed in an Optima XL-I ultracentrifuge, using the interference optics to determine the protein concentration gradient. For each experiment, three cells were assembled with double-sector 12 mm centerpieces and sapphire windows, and loaded with 110 μl of reference buffer (dialysate) and 110 μl of sample, at three different protein concentrations. A blank scan of distilled water was taken before the run, to correct for the effects of window distortion on the fringe displacement data (Yphantis, 1964). The experiments were performed at 4–25°C, at a speed of 13,000 rpm. Fringe displacement data were collected every six to eight hours until equilibrium was reached, as determined by comparison of successive scans using the MATCH program, and the data were edited using the REEDIT program (both programs kindly provided by David Yphantis).

Analysis of sedimentation equilibrium data was performed using the NONLIN program (Johnson *et al.*, 1981). The partial specific volume of KAP-1-RBCC was calculated from the amino acid sequence, and the density of the solvent was calculated according to Laue *et al.* (1992). Three data sets from different loading concen-

trations were fitted simultaneously. Goodness of fit was determined by examination of the residuals and minimization of the variance.

Chemical cross-linking

Chemical cross-linking of recombinant proteins using ethylene glycol bis (succinimidylsuccinate) (EGS) (Pierce) was performed essentially as described (Drees *et al.*, 1997) with the following modifications. For analysis of the oligomerization of wild-type or mutant KAP-1-RBCC proteins, or KOX1-KRAB:KAP-1-RBCC complexes, the final protein concentration was 5 µM in each reaction. The EGS cross-linking was performed at room temperature for five minutes in P300 buffer (pH 7.0), 10% glycerol, 1 mM DTT. The EGS was freshly prepared in dimethylsulfoxide and was added from a stock solution to final concentrations ranging from 0.078 to 5 mM. The reactions were terminated by addition of lysine to a final concentration of 20 mM. After cross-linking, 10 µl of 5× Laemmli sample buffer was added to the reactions, and the samples were analyzed by SDS-PAGE. The proteins were detected by Western blot analysis as described (Ryan *et al.*, 1999).

Expression of KAP-1 and KAP-1-RBCC in baculovirus-infected cells

The KAP-1-RBCC proteins and the full-length KAP-1 protein were expressed individually or co-expressed in baculovirus-infected Sf9 cells (Wistar Core Protein Expression Facility). High-titer baculovirus stocks for each protein were prepared using standard techniques and Sf9 cells were infected for 48 hours. For a 50 ml cell culture, the cell pellets were lysed in 5 ml of 10 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1% NP-40, 1 mM benzamidine, 1 mM PMSF, 2 mM leupeptin, 10 µg/ml pepstatin, 10 µg/ml aprotinin, and 10 mM NaF. The cells were lysed at 4°C for 20 minutes, followed by centrifugation at 16,000 g for ten minutes. The protein was then quantified by Bradford assay (Bio Rad). For protein-protein association assays, the supernatant fraction (250 µg protein) of an individual protein preparation was added to 50 µl of 50% (w/v) Ni-NTA resin slurry, and incubated at 4°C for one hour. The resin was washed with P300 (pH 7.0), 0.1% NP-40, 1 mM PMSF. The bound proteins were extracted from the resin with 5× Laemmli sample buffer, resolved by SDS-PAGE, and detected by Western blot (Ryan *et al.*, 1999).

Acknowledgments

H.P. and D.E.J. were supported by the Wistar Basic Cancer Research Training Grant CA 09171. D.C.S. is supported by DAMD17-98-1-8269. D.W.S. is supported by CA74294 and CA66671. F.J.R. is supported, in part, by National Institutes of Health grants CA 52009, Core grant CA 10815, DK 49210, GM 54220, DAMD 17-96-1-6141, ACS NP-954, the Irving A. Hansen Memorial Foundation, the Mary A. Rumsey Memorial Foundation, and the Pew Scholars Program in the Biomedical Sciences. We thank H.C.M. Nelson for the HSF trimerization domain plasmid, and T.D. Halazonetis for the GCN4-LZ and GCN4-TZ domains plasmid. We thank G. Meroni and A. Ballabio for the MID1 plasmid, and M. Lazar for the TIF1α plasmid. We thank R. Ryan for con-

structing the bv.KAP-1 plasmid, and members of the Speicher, Marmorstein, Berger, and Halazonetis laboratories for many helpful discussions.

References

- Barlow, P. N., Luisi, B., Milner, A., Elliott, M. & Everett, R. (1994). Structure of the C3HC4 domain by 1H-nuclear magnetic resonance spectroscopy. A new structural class of zinc-finger. *J. Mol. Biol.* **237**, 201-211.
- Baxevanis, A. D. & Vinson, C. R. (1993). Interactions of coiled coils in transcription factors: where is the specificity? *Curr. Opin. Genet. Dev.* **3**, 278-285.
- Bellefroid, E. J., Poncelet, D. A., Lecocq, P. J., Revelant, O. & Martial, J. A. (1991). The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *Proc. Natl Acad. Sci. USA*, **88**, 3608-3612.
- Bellini, M., Lacroix, J. C. & Gall, J. G. (1995). A zinc-binding domain is required for targeting the maternal nuclear protein PwA33 to lampbrush chromosome loops. *J. Cell. Biol.* **131**, 563-570.
- Bellon, S. F., Rodgers, K. K., Schatz, D. G., Coleman, J. E. & Steitz, T. A. (1997). Crystal structure of the RAG1 dimerization domain reveals multiple zinc-binding motifs including a novel zinc binuclear cluster. *Nature Struct. Biol.* **4**, 586-591.
- Boddy, M. N., Duprez, E., Borden, K. L. & Freemont, P. S. (1997). Surface residue mutations of the PML RING finger domain alter the formation of nuclear matrix-associated PML bodies. *J. Cell. Sci.* **110**, 2197-2205.
- Borden, K. L. (1998). RING fingers and B-boxes: zinc-binding protein-protein interaction domains. *Biochem. Cell. Biol.* **76**, 351-358.
- Borden, K. L., Boddy, M. N., Lally, J., O'Reilly, N. J., Martin, S., Howe, K., Solomon, E. & Freemont, P. S. (1995a). The solution structure of the RING finger domain from the acute promyelocytic leukaemia proto-oncoprotein PML. *EMBO J.* **14**, 1532-1541.
- Borden, K. L., Lally, J. M., Martin, S. R., O'Reilly, N. J., Etkin, L. D. & Freemont, P. S. (1995b). Novel topology of a zinc-binding domain from a protein involved in regulating early *Xenopus* development. *EMBO J.* **14**, 5947-5956.
- Borden, K. L., Lally, J. M., Martin, S. R., O'Reilly, N. J., Solomon, E. & Freemont, P. S. (1996). In vivo and in vitro characterization of the B1 and B2 zinc-binding domains from the acute promyelocytic leukemia proto-oncoprotein PML. *Proc. Natl Acad. Sci. USA*, **93**, 1601-1606.
- Borden, K. L., Campbell-Dwyer, E. J. & Salvato, M. S. (1997). The promyelocytic leukemia protein PML has a pro-apoptotic activity mediated through its RING domain. *FEBS Letters*, **418**, 30-34.
- Bryan, J. K. (1977). Molecular weights of protein multimers from polyacrylamide gel electrophoresis. *Anal. Biochem.* **78**, 513-519.
- Buchner, G., Montini, E., Andolfi, G., Quaderi, N., Cainarca, S., Messali, S., Bassi, M. T., Ballabio, A., Meroni, G. & Franco, B. (1999). MID2, a homologue of the opitz syndrome gene MID1: similarities in subcellular localization and differences in expression during development. *Hum. Mol. Genet.* **8**, 1397-1407.
- Cainarca, S., Messali, S., Ballabio, A. & Meroni, G. (1999). Functional characterization of the opitz syn-

drome gene product (midin): evidence for homodimerization and association with microtubules throughout the cell cycle. *Hum. Mol. Genet.* **8**, 1387-1396.

Cao, T., Borden, K. L. B., Freemont, P. S. & Etkin, L. D. (1997). Involvement of the rfp tripartite motif in protein-protein interactions and subcellular distribution. *J. Cell Sci.* **110**, 1563-1571.

Cao, T., Duprez, E., Borden, K. L., Freemont, P. S. & Etkin, L. D. (1998). Ret finger protein is a normal component of PML nuclear bodies and interacts directly with PML. *J. Cell Sci.* **111**, 1319-1329.

Drees, B. L., Grotkopp, E. K. & Nelson, H. C. (1997). The GCN4 leucine zipper can functionally substitute for the heat shock transcription factor's trimerization domain. *J. Mol. Biol.* **273**, 61-74.

Dyck, J. A., Maul, G. G., Miller, W. H., Jr, Chen, J. D., Kakizuka, A. & Evans, R. M. (1994). A novel macromolecular structure is a target of the promyelocytic-retinoic acid receptor oncoprotein. *Cell*, **76**, 333-343.

El-Husseini, A. E. & Vincent, S. R. (1999). Cloning and characterization of a novel RING finger protein that interacts with class V myosins. *J. Biol. Chem.* **274**, 19771-19777.

Fagioli, M., Alcalay, M., Tomassoni, L., Ferrucci, P. F., Mencarelli, A., Riganelli, D., Grignani, F., Pozzan, T., Nicoletti, I. & Pelicci, P. G. (1998). Cooperation between the RING + B1-B2 and coiled-coil domains of PML is necessary for its effects on cell survival. *Oncogene*, **16**, 2905-2913.

Fredericks, W. J., Galili, N., Mukhopadhyay, S., Rovera, G., Bennicelli, J., Barr, F. G. & Rauscher, F., Jr (1995). The PAX3-FKHR fusion protein created by the t(2;13) translocation in alveolar rhabdomyosarcomas is a more potent transcriptional activator than PAX3. *Mol. Cell. Biol.* **15**, 1522-1535.

Friedman, J. R., Fredericks, W. J., Jensen, D. E., Speicher, D. W., Huang, X. P., Neilson, E. G. & Rauscher, F., Jr (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev.* **10**, 2067-2078.

Goddard, A. D., Borrow, J., Freemont, P. S. & Solomon, E. (1991). Characterization of a zinc finger gene disrupted by the t(15;17) in acute promyelocytic leukemia. *Science*, **254**, 1371-1374.

Grignani, F., Testa, U., Roggia, D., Ferrucci, P. F., Samoggia, P., Pinto, A., Aldinucci, D., Gelmetti, V., Fagioli, M., Alcalay, M., Seeler, J., Nicoletti, I., Peschle, C. & Pelicci, P. G. (1996). Effects on differentiation by the promyelocytic leukemia PML/RARalpha protein depend on the fusion of the PML protein dimerization and RARalpha DNA binding domains. *EMBO J.* **15**, 4949-4958.

Hasegawa, N., Iwashita, T., Asai, N., Murakami, H., Iwata, Y., Isomura, T., Goto, H., Hayakawa, T. & Takahashi, M. (1996). A RING finger motif regulates transforming activity of the rfp/ret fusion gene. *Biochem. Biophys. Res. Commun.* **225**, 627-631.

Hemenway, C. S., Halligan, B. W. & Levy, L. S. (1998). The Bmi-1 oncprotein interacts with dinG and MPH2: the role of RING finger domains. *Oncogene*, **16**, 2541-2547.

Hope, I. A. & Struhl, K. (1987). GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. *EMBO J.* **6**, 2781-2784.

Hurst, H. C. (1995). Transcription factors 1: bZIP proteins. *Protein Profile*, **2**, 101-168.

Irminger-Finger, I., Siegel, B. D. & Leung, W. C. (1999). The functions of breast cancer susceptibility gene 1 (BRCA1) product and its associated proteins. *Biol. Chem.* **380**, 117-128.

Johnson, M. L., Correia, J. J., Yphantis, D. A. & Halvorson, H. R. (1981). Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.* **36**, 575-588.

Kakizuka, A., Miller, W. H., Jr, Umesono, K., Warrell, R. P., Jr, Frankel, S. R., Murty, V. V., Dmitrovsky, E. & Evans, R. M. (1991). Chromosomal translocation t(15;17) in human acute promyelocytic leukemia fuses RAR alpha with a novel putative transcription factor, PML. *Cell*, **66**, 663-674.

Laue, T. M., Shah, B., Ridgeway, T. M. & Pelletier, S. L. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J. & Horton, L. C., eds), pp. 90-125, Royal Society of Chemistry, Cambridge.

Le Douarin, B., Zechel, C., Garnier, J. M., Lutz, Y., Tora, L., Pierrat, P., Heery, D., Gronemeyer, H., Chambon, P. & Losson, R. (1995). The N-terminal part of TIF1, a putative mediator of the ligand-dependent activation function (AF-2) of nuclear receptors, is fused to B-raf in the oncogenic protein T18. *EMBO J.* **14**, 2020-2033.

Lupas, A. (1996). Coiled coils: new structures and new functions. *Trends Biochem. Sci.* **21**, 375-382.

Margolin, J. F., Friedman, J. R., Meyer, W. K., Vissing, H., Theisen, H. J. & Rauscher, F., Jr (1994). Kruppel-associated boxes are potent transcriptional repression domains. *Proc. Natl Acad. Sci. USA*, **91**, 4509-4513.

Maul, G. G. (1998). Nuclear domain 10, the site of DNA virus transcription and replication. *Bioessays*, **20**, 660-667.

McWhirter, S. M., Pullen, S. S., Holton, J. M., Crute, J. J., Kehry, M. R. & Alber, T. (1999). Crystallographic analysis of CD40 recognition and signaling by human TRAF2. *Proc. Natl Acad. Sci. USA*, **96**, 8408-8413.

Meza, J. E., Brzovic, P. S., King, M. C. & Klevit, R. E. (1999). Mapping the functional domains of BRCA1. Interaction of the ring finger domains of BRCA1 and BARD1. *J. Biol. Chem.* **274**, 5659-5665.

Miki, T., Fleming, T. P., Crescenzi, M., Molloy, C. J., Blam, S. B., Reynolds, S. H. & Aaronson, S. A. (1991). Development of a highly efficient expression cDNA cloning system: application to oncogene isolation. *Proc. Natl Acad. Sci. USA*, **88**, 5167-5171.

Moosmann, P., Georgiev, O., Le, Douarin, B., Bourquin, J. P. & Schaffner, W. (1996). Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. *Nucl. Acids Res.* **24**, 4859-4867.

Ogawa, S., Goto, W., Oriomo, A., Hosoi, T., Ouchi, Y., Muramatsu, M. & Inoue, S. (1998). Molecular cloning of a novel RING finger-B box-coiled coil (RBCC) protein, terf, expressed in the testis. *Biochem. Biophys. Res. Commun.* **251**, 515-519.

Park, Y. C., Burkitt, V., Villa, A. R., Tong, L. & Wu, H. (1999). Structural basis for self-association and receptor recognition of human TRAF2. *Nature*, **398**, 533-538.

Pengue, G., Caputo, A., Rossi, C., Barbanti-Brodano, G. & Lania, L. (1995). Transcriptional silencing of human immunodeficiency virus type 1 long terminal repeat-driven gene expression by the Kruppel-associated box repressor domain targeted to the

transactivating response element. *J. Virol.* **69**, 6577-6580.

Perez, A., Kastner, P., Sethi, S., Lutz, Y., Reibel, C. & Chambon, P. (1993). PMLRAR homodimers: distinct DNA binding properties and heteromeric interactions with RXR. *EMBO J.* **12**, 3171-3182.

Perrin, K. & Lacroix, J. C. (1998). XL43 and XL75: two novel RING finger-containing genes expressed during oogenesis and embryogenesis in *Xenopus laevis*. *Gene*, **210**, 127-134.

Pullen, S. S., Labadia, M. E., Ingraham, R. H., McWhirter, S. M., Everdeen, D. S., Alber, T., Crute, J. J. & Kehry, M. R. (1999). High-affinity interactions of tumor necrosis factor receptor-associated factors (TRAFs) and CD40 require TRAF trimerization and CD40 multimerization. *Biochemistry*, **38**, 10168-10177.

Quaderi, N. A., Schweiger, S., Gaudenz, K., Franco, B., Rugarli, E. I., Berger, W., Feldman, G. J., Volta, M., Andolfi, G., Gilgenkrantz, S., Marion, R. W., Hennekam, R. C., Opitz, J. M., Muenke, M., Ropers, H. H. & Ballabio, A. (1997). Opitz G/BBB syndrome, a defect of midline development, is due to mutations in a new RING finger gene on Xp22. *Nature Genet.* **17**, 285-291.

Reddy, B. A. & Etkin, L. D. (1991). A unique bipartite cysteine-histidine motif defines a subfamily of potential zinc-finger proteins. *Nucl. Acids Res.* **19**, 6330.

Ryan, R. F., Schultz, D. C., Ayyanathan, K., Singh, P. B., Friedman, J. R., Fredericks, W. J. & Rauscher, F. J., III (1999). KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol. Cell. Biol.* **19**, 4366-4378.

Saurin, A. J., Borden, K. L., Boddy, M. N. & Freemont, P. S. (1996). Does this have a familiar RING? *Trends Biochem. Sci.* **21**, 208-214.

Shi, P. Y., Maizels, N. & Weiner, A. M. (1997). Recovery of soluble active recombinant protein from inclusion bodies. *Biotechniques*, **23**, 1036-1038.

Sternsdorf, T., Grotzinger, T., Jensen, K. & Will, H. (1997). Nuclear dots: actors on many stages. *Immunobiology*, **198**, 307-331.

Takahashi, M., Ritz, J. & Cooper, G. M. (1985). Activation of a novel human transforming gene, ret, by DNA rearrangement. *Cell*, **42**, 581-588.

Tatematsu, K., Tokunaga, C., Nakagawa, N., Tanizawa, K., Kuroda, S. & Kikkawa, U. (1998). Transcriptional activity of RBCK1 protein (RBCC protein interacting with PKC 1): requirement of RING-finger and B-Box motifs and regulation by protein kinases. *Biochem. Biophys. Res. Commun.* **247**, 392-396.

Venturini, L., You, J., Stadler, M., Galien, R., Lallemand, V., Koken, M. H., Mattei, M. G., Ganser, A., Chambon, P., Losson, R. & de The, H. (1999). TIF1-gamma, a novel member of the transcriptional intermediary factor 1 family. *Oncogene*, **18**, 1209-1217.

Vissing, H., Meyer, W. K., Aagaard, L., Tommerup, N. & Thiesen, H. J. (1995). Repression of transcriptional activity by heterologous KRAB domains present in zinc finger proteins. *FEBS Letters*, **369**, 153-157.

Waterman, M. J., Waterman, J. L. & Halazonetis, T. D. (1996). An engineered four-stranded coiled coil substitutes for the tetramerization domain of wild-type p53 and alleviates transdominant inhibition by tumor-derived p53 mutants. *Cancer Res.* **56**, 158-163.

Weis, K., Rambaud, S., Lavau, C., Jansen, J., Carvalho, T., Carmo-Fonseca, M., Lamond, A. & Dejean, A. (1994). Retinoic acid regulates aberrant nuclear localization of PML-RAR alpha in acute promyelocytic leukemia cells. *Cell*, **76**, 345-356.

Witzgall, R., O'Leary, E., Leaf, A., Onaldi, D. & Bonventre, J. V. (1994). The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *Proc. Natl. Acad. Sci. USA*, **91**, 4514-4518.

Wolf, E., Kim, P. S. & Berger, B. (1997). MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci.* **6**, 1179-1189.

Yphantis, D. A. (1964). Equilibrium ultracentrifugation of dilute solutions. *Biochemistry*, **3**, 297-317.

Edited by P. E. Wright

(Received 22 September 1999; received in revised form 17 November 1999; accepted 18 November 1999)

Targeting Histone Deacetylase Complexes via KRAB-Zinc Finger Proteins: The PHD and Bromodomains of KAP-1 form a Cooperative Repression Domain that Recruits the Mi-2 α Subunit of the NuRD complex

David C. Schultz, Josh R. Friedman, Frank J. Rauscher, III¹

Running title: Histone Deacetylase and KRAB-ZFP's

The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104

¹Corresponding Author

The Wistar Institute
3601 Spruce St.
Philadelphia, PA 19104
Phone: 215-898-0995
Fax: 215-898-3929
Email: Rauscher@wistar.upenn.edu

Abstract

The PHD finger and bromodomain are two highly conserved protein motifs that are spatially linked in a growing family of transcriptional regulatory proteins. Optimal transcriptional repression by KAP-1, a universal corepressor for the KRAB-ZFP superfamily of transcriptional repressors, requires the PHD finger and bromodomain at the COOH-terminus, which form an integrated functional unit. Mutations in the KAP-1 PHD finger that mimic naturally occurring mutations in the PHD finger of hATRX or site-directed mutants in the bromodomain, based on the structure, specifically disrupted the repression function of this domain. Substitution of highly related PHD fingers and bromodomains of other transcriptional regulators failed to functionally complement this repression activity. A yeast two-hybrid screen revealed a specific interaction between this bipartite domain of KAP-1 and the Mi-2 α subunit of the HDAC complex NuRD. These data suggest that the PHD finger and bromodomain function cooperatively as an integrated functional unit that provides an interface for protein-protein interactions. Moreover, the KRAB-ZFP:KAP-1 repression complex may function to target the histone deacetylase and chromatin remodeling activities of the recently defined NuRD complex to specific gene promoters *in vivo*.

Introduction

Regulation of RNA polymerase II involves a complex interplay between DNA-protein interactions and protein-protein interactions. While the general transcription factors regulate the accurate initiation of transcription, it is the proteins that bind gene-specific DNA elements that are instrumental in either negatively or positively regulating the rate of transcription (Hampsey and Reinberg, 1999; Orphanides et al., 1996). Studies aimed at understanding the mechanisms of transcriptional repression have been greatly aided by the realization that the domains which mediate repression are often highly conserved amino-acid sequence motifs which occur in one or more families of proteins with common DNA binding domains.

The KRAB (Krüppel associated box) domain is one example of an abundant amino acid sequence motif found at the NH₂-terminus of nearly one-third of all Krüppel/TFIIIA type C2H₂ zinc finger proteins (Bellefroid et al., 1991). This highly conserved domain displays potent, DNA-binding dependent repression of transcription that requires the KAP-1 corepressor (Friedman et al., 1996; Kim et al., 1996; Margolin et al., 1994; Moosmann et al., 1996; Moosmann et al., 1997). The primary amino acid sequence of KAP-1 revealed the presence of several conserved signature motifs, including a RING finger, B boxes, a coiled-coil region (RBCC), which collectively form an integrated domain that is both necessary and sufficient to directly interact with the KRAB domain (Peng et al., 2000; Peng et al., 2000). The COOH-terminal sequence of KAP-1 revealed a conserved PHD finger and bromodomain. This particular spatial arrangement of motifs has defined an emerging family of transcriptional regulators which include TIF1 α and TIF1 γ (Le Douarin et al., 1995; Venturini et al., 1999). All members of this protein family have been shown to repress transcription when tethered to DNA, and the mechanisms by which they repress transcription are currently being defined. In this

aspect a direct interaction between KAP-1 and mammalian members of the non-histone chromosomal protein, heterochromatin protein 1 (HP1) has been described (Le Douarin et al., 1996; Nielsen et al., 1999; Ryan et al., 1999). Biochemical studies indicate that this interaction is dependent upon the chromoshadow domain of HP1 proteins and a core, hydrophobic pentapeptide sequence (PxVxL) in KAP-1 (Lechner et al., 2000). Furthermore, indirect immunofluorescence studies of interphase nuclei indicated that KAP-1 colocalizes with both heterochromatic and euchromatic HP1 proteins (Ryan et al., 1999). These particular studies suggest that the KRAB:KAP-1 complex may mediate repression in part via the maintenance or initiation of heterochromatic chromosomal environments.

The PHD finger and bromodomain are two highly conserved protein motifs found in an increasing number of proteins with roles in regulating transcription via modification of chromatin structure. Furthermore, these two motifs are commonly linked architecturally in a number of proteins (Bochar et al., 2000; Friedman et al., 1996; Jones et al., 2000; Jones et al., 2000; Le Douarin et al., 1996; Le Douarin et al., 1995; Venturini et al., 1999). Although both motifs are abundant, the function of these two motifs in the regulation of transcription remains to be fully described. The PHD finger (Plant Homeodomain) is a relatively small domain of approximately 60 amino acids defined by a consensus amino acid signature of C4HC3, which is presumed to chelate zinc (Aasland et al., 1995). The biological importance of the PHD finger in protein function is underscored by the identification of naturally occurring germline mutations resulting in single amino acid substitutions in the PHD finger of the hATRX protein, which predispose individuals to α -thalassemia and a variety of developmental defects (Gibbons et al., 1997).

The bromodomain is another conserved signature motif of approximately 100 amino acids that is found in 30-40 independent proteins (Jeanmougin et al., 1997). The solution structure of the bromodomain of pCAF indicates that these amino acids fold into four amphipathic α -helices, and that acetyl-lysine groups of histone tails may be a natural ligand for this structural motif (Dhalluin et al., 1999). In this aspect, synthetically derived mutations in the bromodomain of the yeast proteins RSC1 and GCN5 were observed to impair the chromatin related functions of these proteins *in vitro* and *in vivo*, without effecting the ability of these proteins to assemble into multiprotein complexes (Cairns et al., 1999; Sterner et al., 1999; Syntichaki et al., 2000). These data suggest that the bromodomain is involved in modulating the activity of these transcriptional regulatory complexes.

Our data from the analysis of the PHD finger and bromodomain of KAP-1 suggests that the spatial conservation of these motifs is essential for optimal protein function. These data are consistent with the PHD finger and bromodomain providing an interface for protein-protein interactions, and therefore may have broad implications in understanding the function of proteins that maintain this particular spatial arrangement of motifs. Comparative analysis of proteins containing these two motifs indicates that this bipartite domain is not functionally equivalent among all proteins. Thus, it is likely that the function of this integrated domain will be dependent upon those proteins with which it associates. The interaction of KAP-1 with Mi-2 α /CHD3, is consistent with this hypothesis and implies a role for histone deacetylases in transcriptional repression mediated by the KRAB-ZFP:KAP-1 repressor/corepressor system.

Results

Several pieces of evidence indicate that KAP-1 functions as a transcriptional corepressor for the KRAB repression module, including an intrinsic ability to repress transcription when directly tethered to DNA templates (Agata et al., 1999; Friedman et al., 1996; Moosmann et al., 1996; Nielsen et al., 1999; Ryan et al., 1999). To identify regions of KAP-1 which are required for its transcriptional repression function we have engineered a series of KAP-1 deletion mutants with the heterologous GAL4 DNA binding domain to grossly map repression domains (**Figure 1**). Each heterologous fusion protein was tested for stable expression in COS1 cells and nuclear localization (data not shown). In general, we have found that amino acids COOH-terminal to the RBCC region define at least two independent repression domains within KAP-1 that appear to be additive in nature in order to obtain maximal levels of transcriptional repression. One of these domains includes the recently defined HP1 binding domain of KAP-1 (Lechner et al., 2000; Ryan et al., 1999). Optimal repression of transcription was dependent upon the PHD finger and bromodomain. When autonomously tethered to DNA, these two motifs demonstrated a significant level of repression, suggesting that they can function as an independent repression domain (**Figure 1**). This repression activity was DNA-binding dependent, dose dependent, and universal for all promoters tested, including the SV40 immediate early enhancer, the adenovirus major late promoter, and a minimal genomic promoter fragment of the DNA polymerase β gene (data not shown). Neither the PHD finger nor the bromodomain could independently recapitulate this repression activity, indicating that these two motifs function cooperatively to repress transcription (**Figure 1**). Furthermore, expression of the segment encoding the PHD finger and bromodomain of KAP-1 relieved both GAL4-KRAB and GAL4-KAP-1 directed repression, suggesting that repression by this bipartite domain is mediated by at least one or more titratable

effector molecules (data not shown). The combination of these data suggests that the PHD finger and bromodomain of KAP-1 function as an integrated unit to facilitate repression of transcription by KAP-1. Moreover, these data implicate an additional repression mechanism that is independent of the recruitment of HP1 proteins (Lechner et al., 2000; Ryan et al., 1999).

Mutational analysis of the KAP-1 PHD finger and bromodomain in repression

In order to correlate the molecular relationship between amino acids in the PHD finger and bromodomain and transcriptional repression, we have used a site-directed mutagenesis approach. All amino acid substitutions were made in the context of the heterologous GAL4-PHD finger/bromodomain fusion protein. Stable expression and nuclear localization of all heterologous fusion proteins was confirmed in COS1 cells (**Figure 2C** and data not shown). The effect of each mutation on the repression properties of this bipartite domain was determined in DNA-template based assays. As an initial strategy, we have targeted amino acid residues in the PHD finger of KAP-1 (C628R, V630S, C651F, P654C, and Q657E; **Figure 2A**) which spatially parallel the position of naturally occurring mutations in hATRX that confer an inherited susceptibility to developmental defects (Gibbons et al., 1997). As illustrated in **Figure 2C**, the C628R and C651F mutations significantly impaired the repression activity of this domain, while the P654C mutation demonstrated an intermediate effect. Each of these three amino acids is highly conserved in the core consensus sequence (**Figure 2A**) and may be involved in either metal binding or maintenance of the hydrophobic core. The V630S and Q657E mutations had minimal effects on repression, and therefore, may represent residues that are more variable to the structure and function of the KAP-1 PHD finger.

Second, we targeted amino acid residues that are strictly conserved among PHD fingers most closely related to KAP-1. These data indicate that mutation of any conserved cysteine residue (CC 628, 631 AA, C643A, C666A; **Figure 2A**) nearly abolished all repression by this domain (**Figure 2C**). Mutation of the conserved Trp664 residue yielded a similar result as the cysteine mutations, suggesting that it may have a role in maintaining the structure of this domain. Although the four remaining amino acids mutated (G635A, D636A, H648C, and P660A) are highly conserved among many members of the PHD finger family (**Figure 2A**), mutation of each demonstrated marginal effects on the repression activity of this bipartite domain (**Figure 2C**).

The final scheme focused on the PHD finger involved mutating amino acids in the variable regions of the KAP-1 PHD finger to residues present in the PHD fingers of TIF1 α /TIF1 γ (**Figure 2A**). Only the M639C mutation demonstrated an effect on repression that was comparable to the P654C mutation (**Figure 2C**). The three remaining mutations (C646V, F647Y, and SL 667,668 TF) demonstrated a minimal effect on the repression activity of this domain, suggesting that KAP-1/TIF1 family may utilize the PHD finger and bromodomain to repress transcription through a common biochemical mechanism.

We also used the information provided by the structure of the pCAF bromodomain to design six independent mutations in each of the four α -helices and one mutation in the variable loop region between helix Z and helix A of the KAP-1 bromodomain (**Figure 2B**). As illustrated in **Figure 2C**, mutations in helix B (F761A) and helix C (DD 778,781AA, trunc 781, and F791E) significantly impaired transcriptional repression (**Figure 2C**). The S730I mutation in the variable loop region between helix Z and helix A also displayed an intermediate effect on the repression potential of this domain, reaffirming the potential role for this region in macromolecular recognition. Overall, the data from this mutational analysis is consistent with

the PHD finger and bromodomain of KAP-1 functioning cooperatively, as independent mutations in either domain can significantly disrupt the intrinsic repression potential of this bipartite domain. One interpretation of this data is that these two motifs provide an interface for protein-protein interactions with downstream effectors of transcriptional repression.

Comparative analysis of the KAP-1 PHD finger and bromodomain with domains from related proteins

Mutational analysis of the KAP-1 PHD finger in which mutations designed to make it more similar to its orthologues, TIF1 α and TIF1 γ , suggested that these domains may function via a common biochemical mechanism. To directly test this hypothesis, we have constructed heterologous fusions between the GAL4 DNA binding domain and the PHD finger and bromodomains of TIF1 α , TIF1 γ , and WCRF180 (Bochar et al., 2000) (**Figure 3A**). In addition, we engineered chimeric proteins in which either the PHD finger or bromodomain of TIF1 α , TIF1 γ , hATRX, Mi-2 α , WCRF180, or hGCN5 were substituted for the KAP-1 sequences in order to address the specificity of the repression activity displayed by this bipartite domain (**Figure 3A**). Stable expression and nuclear localization of each chimeric fusion protein was confirmed in COS1 cells (**Figure 3B** and data not shown). When tethered to DNA, the PHD finger and bromodomain of TIF1 α demonstrated a potential to activate transcription of the reporter template, while the PHD finger and bromodomain of TIF1 γ and WCRF180 each failed to demonstrate any appreciable potentiation of transcription (**Figure 3A**). Heterologous substitution of the KAP-1 PHD finger or bromodomain with similar sequences from other transcriptional regulators failed to completely complement the repression activity of the wild-type KAP-1 PHD finger and bromodomain (**Figure 3A**). Interestingly, when the bromodomain

of KAP-1 was tethered with the PHD finger of TIF1 α , the chimeric protein functioned to repress transcription from the reporter template. In general, the KAP-1 bromodomain, when fused to any PHD finger, slightly enhanced the repression potential of the chimeric protein, suggesting that the bromodomain of KAP-1 is essential in mediating specific interactions with effectors which promote repression of transcription (**Figure 3B**). The combination of these data provides additional support that the PHD finger and bromodomain of KAP-1 function as an integrated functional unit to specifically repress transcription. The data also imply that not all PHD fingers and bromodomains are functionally equivalent, despite demonstrating homology in the primary amino acid sequence. Furthermore, the spatial conservation of these two motifs architecturally in proteins is not sufficient to infer a common function for this bipartite domain.

Identification of downstream effectors of KAP-1 PHD/bromodomain-mediated transcriptional repression

To identify potential mechanisms of transcriptional repression directed by the KAP-1 PHD finger and bromodomain, a heterologous fusion between the LEXA DNA binding domain and the PHD finger and bromodomain of KAP-1 was used in a yeast two-hybrid screen. To confirm the specificity of any potential protein interactions identified in this screen, we utilized several of the amino acid substitutions engineered for the evaluation of the structure/function analysis of this bipartite domain. Appropriate expression of the heterologous fusion proteins was confirmed by Western blot analysis and bait proteins displayed no ability to self-activate the selection promoters (data not shown).

From 40 million library transformants four different gene products were recovered based on coactivation of both the integrated LEXA responsive *His3* and *LacZ* reporter genes. The

nucleotide sequence of three, independent transformants of one gene product indicated that each clone possessed an identical fusion junction to the GAL4 activation domain and that the sequence was identical to the 3' nucleotide sequence of the dermatomyositis-specific autoantigen, Mi-2 α /CHD3. Upon reintroduction into yeast, this GAL4 activation domain fusion protein displayed robust ability to activate the *LacZ* reporter with the wild-type LEXA-KAP-1 fusion protein (**Figure 4A**), but failed to activate the reporter with other irrelevant test baits (data not shown). Two PHD finger mutants (V630S and G635A) that demonstrated nearly wild-type levels of transcriptional repression activity also displayed the ability to activate the *LacZ* reporter (**Figure 4A**). In contrast, PHD finger mutants CC628,631AA , C651F, and W664A, which possess impaired repression activity, either failed to activate these reporters or activated with reduced efficiency, respectively (**Figure 4A**). This putative protein-protein interaction was also observed to be dependent on an intact bromodomain, as the 781 truncation mutation displayed reduced affinity for the Mi-2 α in the two-hybrid assay (**Figure 4A**). These observations are completely consistent with the transcriptional effects observed in transient transfection assays (**Figure 2C**) in which the repression activity was dependent upon both domains. Furthermore, mutations in either the PHD finger or the bromodomain were independently sufficient to ablate the repression function of this bipartite domain. The combination of these data strengthens the argument that the PHD finger and bromodomain of KAP-1 function as an integrated functional unit, which provides an interface for protein-protein interactions with molecules that facilitate repression by the KRAB:KAP-1 complex.

Mi-2 α /CHD3 is a member of the CHD family of proteins which were named based on the possession of several signature motifs, Chromodomains, ATPase/Helicases, and DNA binding modules (Delmas et al., 1993; Woodage et al., 1997). Mi-2 α /CHD3 is closely related to

the Mi-2 β /CHD4 protein. Comparison of the amino acids in the putative KAP-1 interaction domain of Mi-2 α /CHD3 to Mi-2 β /CHD4 revealed that these two proteins are 80% identical, yet all three KAP-1 interacting clones rescued in this screen possessed sequence specific to Mi-2 α /CHD3. Based on the sequences of these clones, amino acids 1686 to 2000 harbor the minimal KAP-1 interaction domain (KID) (**Figure 4B**). This sequence encodes a protein with an additional 30 amino acids at the COOH-terminus of the protein that are not present in the putative protein described by Woodage *et al.* for CHD3 (accession number AF006515), yet the remainder of the two sequences were identical. To test the functional significance of the novel 30 amino acids contained in our rescued clones versus the COOH-terminal sequence reported for CHD3, and the ability of Mi-2 β /CHD4 to interact with KAP-1, we engineered fusions of these COOH-terminal sequences with the GAL4 activation domain, respectively (**Figure 4B**). Upon introduction into yeast, the wild-type LEXA-PHD/bromodomain fusion of KAP-1 displayed no activation of the *LacZ* reporter for either protein, indicating that the COOH-termini of these two proteins, CHD3 and Mi-2 β /CHD4, are not functionally equivalent to the Mi-2 α sequence we rescued in the two-hybrid screen (**Figure 4B**). Closer inspection of the COOH-terminal amino acid residues revealed that Mi-2 α /CHD3 quickly diverges from Mi-2 β /CHD4 starting at amino acid 1909 and continues for an additional 91 novel amino acids. It is these amino acid residues that most likely confer the specificity of the Mi-2 α /CHD3 interaction with KAP-1.

In order to evaluate the role of Mi-2 α /CHD3 in KAP-1 mediated transcriptional repression, we chose to disrupt the endogenous KAP-1/Mi-2 α interaction by expressing a dominant negative protein corresponding to amino acids 1738 to 2000 of Mi-2 α (**Figure 5A**). Expression of these amino acids dominantly inhibited KRAB-PAX3 mediated repression of a DNA template in a dose dependent manner (**Figure 5A**). The same effect was observed for

PAX3-KAP-1 directed repression (data not shown). Expression of the dominant negative protein had little effect on the basal activity of the minimal TK promoter, the activity obtained in combination with other non-KRAB PAX3 repression domain fusions, or a PAX3 fusion comprised of the FKHR activation domain (**Figure 5A**, data not shown). Consistent with the yeast two-hybrid data, truncation of the COOH-terminal 30 amino acids of the Mi-2 α KID impaired this inhibitory activity (data not shown). Since Mi-2 α has been described as integral component of a high molecular weight multiprotein complex containing histone deacetylase activity, we tested whether the addition of trichostatin A (TSA) could reverse the repression activity of the KAP-1 PHD finger and bromodomain. **Figure 5B** illustrates that addition of 300 nM TSA partially reversed KAP-1 PHD/bromodomain-mediated repression of three independent promoters. These data provide evidence that Mi-2 α and histone deacetylases may be one potential downstream effector of the KRAB-ZFP:KAP-1 repressor:corepressor complex *in vivo*.

Discussion

The recurrent theme that has emerged from studies aimed at understanding the functional relationships of protein-protein interactions is that highly conserved amino acid signature motifs often function as independent globular domains involved in macromolecular recognition. The identification of these signature amino acid sequence motifs and their spatial organization within a novel protein is often the first clue to its biochemical function. We and others have previously identified a universal corepressor, KAP-1, which binds to the KRAB repression module (Friedman et al., 1996; Kim et al., 1996; Moosmann et al., 1996). Structurally, KAP-1 possesses consensus amino acid sequences for a RING finger, B-boxes, leucine zipper coiled-coil region, PHD finger, and bromodomain (Friedman et al., 1996; Kim et al., 1996; Le Douarin

et al., 1996; Moosmann et al., 1996). Biochemical studies of KAP-1 indicate that the tripartite RBCC region functions as an integrated structural unit that is necessary and sufficient for KRAB recognition and KAP-1 oligomerization (Peng et al., 2000). Mechanistically, a direct interaction with members of the HP1 protein family suggests that KAP-1 may function to repress transcription via the organization of higher order chromatin structure (Le Douarin et al., 1996; Lechner et al., 2000; Nielsen et al., 1999; Ryan et al., 1999). In this report, we present data that builds upon our understanding of KAP-1's function in repression of transcription. We describe a novel transcriptional repression activity defined by the PHD finger and bromodomain of KAP-1. These domains are collectively required for optimal repression of DNA templates by KAP-1, and our studies suggest that these two motifs form a functionally integrated unit, as mutations or heterologous substitutions in either domain disrupt transcriptional repression. These data are consistent with this bipartite domain providing a specific interface for protein-protein interactions. In this regard, we show evidence that the PHD finger and bromodomain of KAP-1 may function to target the histone deacetylase complex, NuRD, to promoter sequences regulated by KRAB-zinc finger proteins.

The PHD finger and bromodomain are two well conserved signature motifs widely distributed among nuclear proteins with established roles in the formation and maintenance of chromatin structure (Aasland et al., 1995; Jeanmougin et al., 1997; Winston and Allis, 1999). Thus, the presence of these two motifs suggested that KAP-1 represses transcription via some chromatin mediated mechanism. Furthermore, the conservation of the PHD finger and bromodomain architecturally in at least three independent families of proteins, implies that the two domains biochemically function together (Bochar et al., 2000; Friedman et al., 1996; Jones et al., 2000; Jones et al., 2000; Le Douarin et al., 1996; Le Douarin et al., 1995; Venturini et al.,

1999). In support of this hypothesis, optimal transcriptional repression by KAP-1 is dependent upon the integrity of this bipartite domain. Comparative analysis of the PHD finger and bromodomain of TIF1 α , TIF1 γ and WCRF180 suggests that the function of this bipartite domain is not equivalent among all proteins that share this particular spatial arrangement of motifs. Heterologous substitution of either of these domains in KAP-1 with related sequences essentially failed to complement the intrinsic transcriptional repression activity of the natural KAP-1 sequence. Thus, it appears from these results that the amino acid sequence of the KAP-1 PHD finger and bromodomain possess an inherent capacity to fold into an integrated functional unit, which may provide a specific interface for protein-protein interactions with downstream effectors of transcriptional repression.

The PHD finger is a small domain characteristically defined by seven cysteines and a histidine that are spatially arranged in a consensus of C4HC3 of varying lengths and composition (Aasland et al., 1995). This particular arrangement of amino acids is highly homologous to the RING finger motif, an interface for protein-protein interactions which is created by chelating two zinc ions in a unique cross-braced fashion (Borden, 2000; Borden, 1998; Jensen et al., 1998; Wu et al., 1996). The biological importance of the PHD finger to protein function is further underscored by the identification of naturally occurring germline mutations which encode for amino acid substitutions in the PHD finger of hATRX that predispose individuals to developmental defects (Gibbons et al., 1997). Our studies represent the first detailed site-directed mutagenesis approach to correlate the molecular anatomy of this domain with a biochemical function. In general, mutation of any highly conserved cysteine residues of the core signature sequence was observed to dramatically effect the repression activity of KAP-1. This observation was not unexpected, as mutations of any of the analogous amino acids in the RING

finger disrupt zinc chelation and protein structure. Mutation of the histidine at the putative fifth metal coordination position to cysteine had marginal effects on the repression activity, but it is possible that the introduced cysteine residue is capable of substituting in the proper coordination of zinc. In addition to these core amino acids which comprise the signature sequence, mutation of W664 significantly impaired the repression activity, while the M639C and P654C mutations demonstrated a moderate effect on repression. The introduction of an extra cysteine residue by the M639C and P654C mutations may have altered proper metal chelation properties of this domain. The W664 amino acid is highly conserved spatially in the related RING finger motif, and mutation of the analogous amino acid in the c-Cbl proto-oncogene yielded a loss of function (Joazeiro et al., 1999). Preliminary structural studies of the PHD finger indicate that this residue is likely to be involved in maintenance of the hydrophobic core of this domain's fold (unpublished data). Overall, the functional significance of any of these mutations will have to wait until a structure for the PHD finger is solved, and/or new biochemical functions for this domain are described. Nonetheless, this panel of mutations will serve as a useful tool in future studies in defining new macromolecular interactions and functions for the PHD finger.

Similar to the PHD finger, the bromodomain is a highly conserved structural motif with a putative role in modulating the activities of multiprotein transcriptional regulatory complexes (Barlev et al., 1998; Cairns et al., 1999; Dhalluin et al., 1999; Ornaghi et al., 1999; Sterner et al., 1999; Syntichaki et al., 2000). Furthermore, the function of the bromodomain in these proteins may require or be influenced by adjacent signature motifs (Cairns et al., 1999; Winston and Allis, 1999). Our data from the analysis of the KAP-1 bromodomain are consistent with these hypotheses. First, it appears that the bromodomain of KAP-1 has intrinsic ability to potentiate repression of transcription. Second, optimal repression of transcription by KAP-1 is dependent

upon the integrity of the bipartite PHD finger and bromodomain, as mutations in either domain are sufficient to significantly impair KAP-1 directed repression of reporter templates. Third, the mutations in the bromodomain that disrupt the repression activity are consistent with this domain's involvement in macromolecular recognition. Finally, significant evidence suggests that some feature of histones or nucleosomal structure serve as a ligand for the bromodomain. The specific interaction between the PHD finger and bromodomain of KAP-1 with the Mi-2 α subunit of the NuRD complex may require some interface provided by the potential interaction between the bromodomain of KAP-1 and nucleosomes. This hypothesis would be consistent with recent genetic evidence that indicates the bromodomain of GCN5 is required for the chromatin remodeling activities of the SWI/SNF complex at a genomic locus in yeast (Syntichaki et al., 2000). In this regard, both the *His3* and the *LacZ* reporters used in the selection process for positive interacting clones in the two-hybrid assay were chromosomally integrated. Furthermore, this potential ternary complex may explain why no polypeptides with an obvious targeting mechanism were identified in the original biochemical purifications of NuRD (Tong et al., 1998; Xue et al., 1998; Zhang et al., 1998). Both *Mi-2 α* and *Mi-2 β* have been describe as integral components of a multi-protein complex containing both histone deacetylase and chromatin remodeling activities (Tong et al., 1998; Wade et al., 1998; Xue et al., 1998; Zhang et al., 1998). The association of KAP-1 with *Mi-2 α /CHD3* builds upon our understanding of how the KRAB-ZFP:KAP-1 repressor:corepressor system mechanistically regulates gene expression. Moreover, the interaction between KAP-1 and *Mi-2 α /CHD3* may be one alternative mechanism by which the biochemical activities of the NuRD complexes are targeted to specific genes *in vivo*.

Several pieces of evidence have implied a role for the acetylation status of histone tails with mechanisms of long term silencing of gene expression. Genetic analyses in *D. melanogaster* have demonstrated a potential role for dMi-2 in polycomb silencing of homeotic genes (Kehle et al., 1998). Consistent with this genetic observation, the sequence specific transcription factor Ikaros was observed to colocalize with components of the NuRD complex and the heterochromatin protein 1(M31), during lymphoid cell development (Kim et al., 1999). In this regard, we and others have reported a direct biochemical interaction between KAP-1 and mammalian homologues of the HP1 protein family, and have shown that KAP-1 and mammalian HP1 proteins can physically occupy the same spatial domains in interphase nuclei of NIH/3T3 cells (Ryan et al., 1999). Extensive genetic analyses of both the polycomb and HP1 protein families have firmly established a role for these proteins in epigenetic gene silencing in *D. melanogaster*. Furthermore, genetic screens for dominant suppressors of position-effect-variegation in *D. melanogaster* have identified mutant alleles of *HDAC1* (Mottus et al., 2000). The combination of these data implies a potential link between multi-protein complexes that promote transcriptional repression by histone deacetylation and repression mediated by organization of high-order chromatin structure. In support of this hypothesis, SAP30, a novel member of the SIN3 histone deacetylase complex, has been observed to effect telomere silencing in yeast (Zhang et al., 1998). This apparent coupling of two independent repression mechanisms may lead to synergistic repression of their apparent target genes.

The relative abundance of the KRAB domain may potentially make it one of the single most common mechanisms of transcriptional repression directed by sequence-specific factors in vertebrates. Transcriptional repression by a KRAB domain requires interaction with the corepressor, KAP-1. Thus, it appears that the effector molecules of KRAB-mediated

transcriptional repression are likely due to a network of protein interactions with KAP-1. We suggest that the KAP-1 corepressor may function to repress transcription through several distinct mechanisms, including histone deacetylation and heterochromatinization. Unlike prokaryotic genes, eukaryotic genes exist in a "restrictive" state by virtue of nucleosome based chromatin structure (Struhl, 1999). In this model, the role of sequence specific repressors like the superfamily of KRAB-ZFP's may be to create a silent gene environment completely resistant to activator and/or core transcription machinery binding. It may be that corepressor molecules, like KAP-1, exist to assist in recruiting independent biochemical activities that function combinatorially to optimally repress a target locus *in vivo*. The abundance of KRAB domain proteins and the capacity of the KAP-1 corepressor to repress transcription through multiple independent mechanisms increase the potential of the KRAB:KAP-1 repressor:corepressor system to serve as a universal regulator of transcription.

Materials and Methods

The pM2-KAP-1 (293-835) mammalian expression vector has been previously described (Friedman et al., 1996). The GAL4-KAP-1 fusion proteins including 408-835, 408-618, 478-835, 478-680, 478-618 were created by PCR using the primers 408for (*Bam*HI) 5'-CCGGGATCCAGATTGTGGCAGAGCGTCCTG-3', 478for(*Bam*HI) 5'-CCGGGATCCAGGTGAGCGGCCTTATGCGC-3', 618rev(*Hind*III) 5'-GATAAGCTTCACGGGCCACCACCTGGGGC-3', 835rev(*Kba*I) 5'-GCTATCTAGACTAAATGGTGGCACTGTCATCCAGG-3', respectively. The resulting PCR products were digested with the restriction endonucleases indicated in brackets preceding the primer sequence and cloned into the corresponding restriction sites of the vector pM1. The GAL4-KAP-1 (619-835) expression plasmid was constructed by subcloning a *Xma*I fragment

from pBL-KAP-1 into the *Xma*I site of pM1. The GAL4-KAP-1 (619-679) expression plasmid was created by PCR using the primers, 619for(*Eco*RI) 5'-TAGCGAATTCTGGAACCTGGATGACAGTGC-3' and 679rev(*Sal*I) 5'-ATCGGTGGACATCCTCCTCCTCAGGTCAG-3'. The resulting PCR product was digested with the restriction endonucleases indicated in brackets and cloned into the corresponding sites of the vector pM1. Carboxy-terminal to codon 679 of KAP-1 the protein reads VDASAEASR*. The GAL4-KAP-1 (674-835) expression plasmid was created by PCR, using the primers, 674for(*Bam*HI) 5'-ATCGGATCCGACCTGAAGGAGGAGGATGGC-3' and 835rev(*Hind*III) 5'-GATCCCGGAAGCTTCAGGGGCCATCACCTGG-3'. The resulting PCR product was digested with the restriction endonucleases indicated in brackets and cloned into the corresponding sites of the vector pM2.

Site-directed point mutations in the KAP-1 PHD finger and bromodomain were engineered by standard overlap extension PCR-mediated mutagenesis procedures. The mutagenic primers for the described mutations contained the following codons: CC628,631AA, TGC to GCC; C628R, TGC to CGC; V630S, GTC to AGC; G635A, GGC to GCC; D636A, GAT to GCT; M639C, ATG to TGC; C643A, TGT to GCT; C646V, TGT to GTT; F647Y, TTC to TAC; H648C, CAC to TGC; C651F, TGT to TTT; P654C, CCG to TGC; Q657E, CAG to GAG; P660A, CCA to GCA; W664A, TGG to GCG; C666A, TGC to GCC; SL667,668TF, TCA and CTC to ACA and TTC; E715K, GAA to AAA; S730I, TCC to ATC; L742T, CTG to ACG; F761A, TTT to GCT; D778,781AA, GAC to GCT; F791E, TTC to GAG. The pM1-KAP-1 (619-835) trunc781 resulted from spontaneous deletion of G2444 from the nucleotide sequence during PCR of the D778,781AA mutation and a reading frame shift that results in a protein with 15 novel amino acids, AKAACSPSSGCSASSRRA* starting at codon 778 of KAP-1.

The GAL4 expression plasmids for the PHD finger and bromodomain of TIF1 α (amino acids 785-1016), TIF1 γ (amino acids 881-1120), and WCRF180 (amino acids 1141-1556) were created by PCR using the primers TIF1 α for(*Sma*I) 5'-
GATCCCGGGAGGAAGGAGGATGACCCC-3', TIF1 α rev(*Sma*I) 5'-
GTACCCGGGTTACTTAAGCAGCTGGCG-3', TIF1 γ for(*Sma*I) 5'-
GATCCCGGGAAATAAAGATGATGACCC-3', TIF1 γ rev(*Sma*I) 5'-
CTGCCCGGGTCATCTGACTTTAGGCGT-3', WCRF180for(*Sma*I) 5'-
GATCCCGGGATATGGTCTAAATCTATA-3', and WCRF180rev(*Sma*I) 5'-
CTGCCCGGGTCAGATT CGTGACTTTGC-3', respectively. The resulting PCR products were digested with the restriction endonucleases indicated in brackets and cloned into the corresponding sites of either pM1 or pM3. The PHD finger and bromodomain chimeras were engineered by standard overlap extension PCR-mediated mutagenesis procedures. Forward primers were: ATRX 5'(*Sma*I) 5'-GACCCGGGTAGCCGTGACTCAGATGG-3', Mi-2 α 5'(*Sma*I) 5'-GATCCCGGGATGGCTACGAGACGGATC-3', and as indicated above for WCRF180 and KAP-1. Fusion primers were as follows: TIF1 α /KAP-1for 5'-
AAGCCAGAGGTTGACTATGGCAGCCTCAGCCTGGAT-3', TIF1 γ /KAP-1for 5'-
AAGCCAGAACGTTGAATATGGCAGCCTCAGCCTGGAT-3', ATRX/KAP-1for 5'-
ACTGCATGTAACAGCGTAGGCAGCCTCAGCCTCAGC-3', Mi-2 α /KAP-1for 5'-
GTCCAGTGGAGGCCAAGGGCAGCCTCAGCCTCAGC-3', WCRF180/KAP-1for 5'-
CGTTCTAGAAGACTCTCCGGCAGCCTCAGCCTCAGC-3', KAP-1/TIF1 α for 5'-
GACCTGAAGGAGGAGGATGATTGTGATGTTCCCAGT-3', KAP-1/TIF1 γ for 5'-
GACCTGAAGGAGGAGGATGATTGTGATAATTGCAA-3', KAP-1/WCRF180for 5'-
GACCTGAAGGAGGAGGATAGATCTGTAAATATTGCT-3', KAP-1/GCN5for 5'

GACCTGAAGGAGGAGGATACAGGCTGGAAGCCATTG-3'. Reverse primers were as follows: GCN5rev(*Sma*I) 5'-GATCCCAGGCTACTTGTCAATGAGGCC-3', and as indicated above for KAP-1, TIF1 α , TIF1 γ , and WCRF180. Second-round PCR products were digested with restriction endonucleases as defined in the parenthesis preceding the primer sequence and cloned into the corresponding sites of either pM1 or pM3.

The LEXA-KAP-1 (amino acids 619-835) yeast expression plasmid was generated by the ligation of a *Xma*I fragment from pBL-KAP-1 into pBTM116. PHD finger and bromodomain mutants were subcloned from pM1 to pBTM116 as a *Sma*I fragment. GAL4 activation domain fusions with the C-terminus of Mi-2 α del and Mi-2 β were created by PCR using the following primers, respectively: Mi-2 α for(*Bam*HI) 5'-

GTAACGGATCCAGGGCCTCGAGATGAGGCCAC-3', Mi-2 α delta CTrev(*Bam*HI) 5'-

GATCGGATCCTCACGTTGGTGGCGGCTGTGATGAAG-3', Mi-2 β for(*Bam*HI) 5'-

GTAACGGATCCAAATGGAGAGACCCCCAAGGACCTG-3', Mi-2 β rev(*Bam*HI) 5'-

GATCGGATCCTCACTGCTGGCTACCTGCTG-3'. Resulting PCR products were digested with *Bam*HI and cloned into the corresponding restriction site of pACTII (Clontech).

The CMV-Mi-2 α ct mammalian expression plasmid was created by subcloning a *Bgl*III/*Bam*HI fragment from the rescued cDNA in-frame with an N-terminal 6His-tag possessing a consensus Kozak sequence and initiator methionine. The KRAB-PAX3 and the PAX3-KAP-1 expression constructs have been described previously (Fredericks et al., 2000). Appropriate reading frame fusions and integrity of flanking sequences for all constructs created by PCR was confirmed by DNA sequence analysis of both strands.

Yeast two-hybrid: The yeast two-hybrid system as modified by Stan Hollenberg was used for all yeast experiments. A human oligo-dT-primed B-cell cDNA library was screened as previously described (Jensen et al., 1998).

Transient transfection: Protein expression from all plasmids was confirmed by transient transfection of COS-1 cells followed by immunoprecipitation of 35 S-methionine labeled cell extracts, as described previously. All transcription assay transfections were done as previously described (Ryan et al., 1999).

Acknowledgments

D.C.S. was supported in part by the Wistar Basic Cancer Research Training Grant CA 09171 and DAMD 17-98-1-8269 grant. We would like to acknowledge Robert F. Ryan for doing the original two-hybrid screen. We thank Trevor Woodage for Mi-2 α cDNA sequences, David Picketts for hATRX PHD finger sequence, and Shelley Berger for hGCN5 bromodomain sequences. F.J.R. is supported in part by National Institutes of Health grants CA 52009, Core grant CA 10815, DK 49210, GM 54220, DAMD 17-96-1-6141, ACS NP-954, the Irving A. Hansen Memorial Foundation, the Mary A. Rumsey Memorial Foundation, and the Pew Scholars Program in the Biomedical Sciences.

References

Aasland, R., Gibson, T. J., and Stewart, A. F. 1995. The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem Sci* 20: 56-59.

Agata, Y., Matsuda, E., and Shimizu, A. 1999. Two novel Kruppel-associated box-containing zinc-finger proteins, KRAZ1 and KRAZ2, repress transcription through functional interaction with the corepressor KAP-1 (TIF1beta/KRIP-1). *J Biol Chem* **274**: 16412-16422.

Barlev, N. A., Poltoratsky, V., Owen-Hughes, T., Ying, C., Liu, L., Workman, J. L., and Berger, S. L. 1998. Repression of GCN5 histone acetyltransferase activity via bromodomain- mediated binding and phosphorylation by the Ku-DNA-dependent protein kinase complex. *Mol Cell Biol* **18**: 1349-1358.

Bellefroid, E. J., Poncelet, D. A., Lecocq, P. J., Revelant, O., and Martial, J. A. 1991. The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *Proc Natl Acad Sci U S A* **88**: 3608-3612.

Bochar, D. A., Savard, J., Wang, W., Lafleur, D. W., Moore, P., Cote, J., and Shiekhattar, R. 2000. A family of chromatin remodeling factors related to williams syndrome transcription factor [In Process Citation]. *Proc Natl Acad Sci U S A* **97**: 1038-1043.

Borden, K. L. 2000. RING Domains: Master Builders of Molecular Scaffolds? *J Mol Biol* **295**: 1103-1112.

Borden, K. L. 1998. RING fingers and B-boxes: zinc-binding protein-protein interaction domains. *Biochem Cell Biol* **76**: 351-358.

Cairns, B. R., Schlichter, A., Erdjument-Bromage, H., Tempst, P., Kornberg, R. D., and Winston, F. 1999. Two functionally distinct forms of the RSC nucleosome-remodeling complex, containing essential AT hook, BAH, and bromodomains. *Mol Cell* **4**: 715-723.

Delmas, V., Stokes, D. G., and Perry, R. P. 1993. A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-like helicase domain. *Proc Natl Acad Sci U S A* **90**: 2414-2418.

Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M. 1999. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**: 491-496.

Fredericks, W. J., Ayyanathan, K., Herlyn, M., Friedman, J. R., and Rauscher, F. J. III. 2000. An engineered PAX3-KRAB transcriptional repressor inhibits the malignant phenotype of aveolar rhabdomyosarcoma cells harboring the endogenous PAX3-FKHR oncogene. *Mol Cell Biol in press*.

Friedman, J. R., Fredericks, W. J., Jensen, D. E., Speicher, D. W., Huang, X. P., Neilson, E. G., and Rauscher, F. J., III 1996. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev* **10**: 2067-2078.

Gibbons, R. J., Bachoo, S., Picketts, D. J., Aftimos, S., Asenbauer, B., Bergoffen, J., Berry, S. A., Dahl, N., Fryer, A., Keppler, K., Kurosawa, K., Levin, M. L., Masuno, M., Neri, G., Pierpont, M. E., Slaney, S. F., and Higgs, D. R. 1997. Mutations in transcriptional regulator ATRX establish the functional significance of a PHD-like domain [letter]. *Nat Genet* **17**: 146-148.

Hampsey, M., and Reinberg, D. 1999. RNA polymerase II as a control panel for multiple coactivator complexes. *Curr Opin Genet Dev* 9: 132-139.

Jeanmougin, F., Wurtz, J. M., Le Douarin, B., Chambon, P., and Losson, R. 1997. The bromodomain revisited [letter]. *Trends Biochem Sci* 22: 151-153.

Jensen, D. E., Proctor, M., Marquis, S. T., Gardner, H. P., Ha, S. I., Chodosh, L. A., Ishov, A. M., Tommerup, N., Vissing, H., Sekido, Y., Minna, J., Borodovsky, A., Schultz, D. C., Wilkinson, K. D., Maul, G. G., Barlev, N., Berger, S. L., Prendergast, G. C., and Rauscher, F. J., III 1998. BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression. *Oncogene* 16: 1097-1112.

Joazeiro, C. A., Wing, S. S., Huang, H., Leverson, J. D., Hunter, T., and Liu, Y. C. 1999. The tyrosine kinase negative regulator c-Cbl as a RING-type, E2- dependent ubiquitin-protein ligase [see comments]. *Science* 286: 309-312.

Jones, M. H., Hamana, N., Nezu, J., and Shimane, M. 2000. A Novel Family of Bromodomain Genes. *Genomics* 63: 40-45.

Jones, M. H., Hamana, N., and Shimane, M. 2000. Identification and Characterization of BPTF, a Novel Bromodomain Transcription Factor. *Genomics* 63: 35-39.

Kehle, J., Beuchle, D., Treuheit, S., Christen, B., Kennison, J. A., Bienz, M., and Muller, J. 1998. dMi-2, a hunchback-interacting protein that functions in polycomb repression. *Science* **282**: 1897-1900.

Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T., Kingston, R., and Georgopoulos, K. 1999. Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. *Immunity* **10**: 345-355.

Kim, S. S., Chen, Y. M., O'Leary, E., Witzgall, R., Vidal, M., and Bonventre, J. V. 1996. A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins. *Proc Natl Acad Sci U S A* **93**: 15299-15304.

Le Douarin, B., Nielsen, A. L., Garnier, J. M., Ichinose, H., Jeanmougin, F., Losson, R., and Chambon, P. 1996. A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. *Embo J* **15**: 6701-6715.

Le Douarin, B., Zechel, C., Garnier, J. M., Lutz, Y., Tora, L., Pierrat, P., Heery, D., Gronemeyer, H., Chambon, P., and Losson, R. 1995. The N-terminal part of TIF1, a putative mediator of the ligand-dependent activation function (AF-2) of nuclear receptors, is fused to B-raf in the oncogenic protein T18. *Embo J* **14**: 2020-2033.

Lechner, M. S., Begg, G. E., Speicher, D. W., and Rauscher, F. J. III. 2000. Molecular determinants for targeting HP1-mediated gene silencing: a direct chromoshadow domain-KAP-1 corepressor interaction is essential. *Mol Cell Biol in press.*

Margolin, J. F., Friedman, J. R., Meyer, W. K., Vissing, H., Thiesen, H. J., and Rauscher, F. J., III 1994. Kruppel-associated boxes are potent transcriptional repression domains. *Proc Natl Acad Sci U S A* **91**: 4509-4513.

Moosmann, P., Georgiev, O., Le Douarin, B., Bourquin, J. P., and Schaffner, W. 1996. Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. *Nucleic Acids Res* **24**: 4859-4867.

Moosmann, P., Georgiev, O., Thiesen, H. J., Hagmann, M., and Schaffner, W. 1997. Silencing of RNA polymerases II and III-dependent transcription by the KRAB protein domain of KOX1, a Kruppel-type zinc finger factor. *Biol Chem* **378**: 669-677.

Mottus, R., Sobel, R. E., and Grigliatti, T. A. 2000. Mutational analysis of a histone deacetylase in *Drosophila melanogaster*: missense mutations suppress gene silencing associated with position effect variegation. *Genetics* **154**: 657-668.

Nielsen, A. L., Ortiz, J. A., You, J., Oulad-Abdelghani, M., Khechumian, R., Gansmuller, A., Chambon, P., and Losson, R. 1999. Interaction with members of the heterochromatin protein 1

(HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. *Embo J* **18**: 6385-6395.

Ornaghi, P., Ballario, P., Lena, A. M., Gonzalez, A., and Filetici, P. 1999. The bromodomain of Gcn5p interacts in vitro with specific residues in the N terminus of histone H4. *J Mol Biol* **287**: 1-7.

Orphanides, G., Lagrange, T., and Reinberg, D. 1996. The general transcription factors of RNA polymerase II. *Genes Dev* **10**: 2657-2683.

Peng, H., Begg, G. E., Harper, S. L., Friedman, J. R., Speicher, D. W., and Rauscher, F. J. III 2000. Biochemical analysis of the KRAB transcriptional repression domain:spectral, kinetic and stoichiometric properties of the KRA-KAP-1 complex. *J Biol Chem in press*.

Peng, H., Begg, G. E., Schultz, D. C., Friedman, J. R., Jensen, D. E., Speicher, D. W., and Rauscher, F. J., III 2000. Reconstitution of the KRAB-KAP-1 Repressor Complex: A Model System for Defining the Molecular Anatomy of RING-B Box-coiled-coil Domain- mediated Protein-protein Interactions. *J Mol Biol* **295**: 1139-1162.

Ryan, R. F., Schultz, D. C., Ayyanathan, K., Singh, P. B., Friedman, J. R., Fredericks, W. J., and Rauscher, F. J., III 1999. KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol Cell Biol* **19**: 4366-4378.

Sterner, D. E., Grant, P. A., Roberts, S. M., Duggan, L. J., Belotserkovskaya, R., Pacella, L. A., Winston, F., Workman, J. L., and Berger, S. L. 1999. Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction. *Mol Cell Biol* **19**: 86-98.

Struhl, K. 1999. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell* **98**: 1-4.

Syntichaki, P., Topalidou, I., and Thireos, G. 2000. The Gcn5 bromodomain co-ordinates nucleosome remodelling. *Nature* **404**: 414-417.

Tong, J. K., Hassig, C. A., Schnitzler, G. R., Kingston, R. E., and Schreiber, S. L. 1998. Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. *Nature* **395**: 917-921.

Venturini, L., You, J., Stadler, M., Galien, R., Lallemand, V., Koken, M. H., Mattei, M. G., Ganser, A., Chambon, P., Losson, R., and de The, H. 1999. TIF1gamma, a novel member of the transcriptional intermediary factor 1 family. *Oncogene* **18**: 1209-1217.

Wade, P. A., Jones, P. L., Vermaak, D., and Wolffe, A. P. 1998. A multiple subunit Mi-2 histone deacetylase from *Xenopus laevis* cofractionates with an associated Snf2 superfamily ATPase. *Curr Biol* **8**: 843-846.

Winston, F., and Allis, C. D. 1999. The bromodomain: a chromatin-targeting module? [news]. *Nat Struct Biol* **6**: 601-604.

Woodage, T., Basrai, M. A., Baxevanis, A. D., Hieter, P., and Collins, F. S. 1997. Characterization of the CHD family of proteins. *Proc Natl Acad Sci U S A* **94**: 11472-11477.

Wu, L. C., Wang, Z. W., Tsan, J. T., Spillman, M. A., Phung, A., Xu, X. L., Yang, M. C., Hwang, L. Y., Bowcock, A. M., and Baer, R. 1996. Identification of a RING protein that can interact in vivo with the BRCA1 gene product. *Nat Genet* **14**: 430-440.

Xue, Y., Wong, J., Moreno, G. T., Young, M. K., Cote, J., and Wang, W. 1998. NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Mol Cell* **2**: 851-861.

Zhang, Y., LeRoy, G., Seelig, H. P., Lane, W. S., and Reinberg, D. 1998. The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell* **95**: 279-289.

Zhang, Y., Sun, Z. W., Iratni, R., Erdjument-Bromage, H., Tempst, P., Hampsey, M., and Reinberg, D. 1998. SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. *Mol Cell* **1**: 1021-1031.

Figure Legends

Figure 1. Analysis of the intrinsic repression activity of KAP-1 indicates that the PHD finger and bromodomain are required and sufficient to repress transcription. (A) Schematic illustration of the heterologous GAL4-KAP-1 constructs, as defined by the amino acid numbers at the left. The intrinsic repressor activity of KAP-1 was measured in a transient assay, using a minimal TK-luciferase reporter template regulated by 5 consensus GAL4 UAS. All experiments were done in NIH/3T3 cells with 0.5 µg of reporter plasmid and 5 µg of the indicated heterologous GAL4-KAP-1 construct. Fold repression represents the ratio of luciferase activity measured for the reporter alone to the activity measured in the presence of the indicated effector proteins after normalization for transfection efficiency. Error bars represent the standard deviation for three independent transfections. Absence of error bars indicates a standard deviation too little to physically illustrate.

Figure 2. Mutations in the PHD finger and bromodomain significantly impair the intrinsic repression activity of this integrated transcriptional repression domain. (A) Amino acid sequence alignment of the KAP-1 PHD finger with related sequences from 13 independent proteins. The numbers in brackets indicate the corresponding amino acids in each protein. Strictly conserved amino acids are shaded in black. Boxed amino acids in which the chemical nature of the side chain has been conserved. Asterisks indicate amino acids that were mutated to match naturally occurring mutations in the hATRX protein. The pound (#) symbol represents conserved amino acids that were mutated. Dollar signs (\$) represent amino acids in KAP-1 which were mutated to match the corresponding amino acid in TIF1 α /TIF1 γ . Each mutation was made in the context

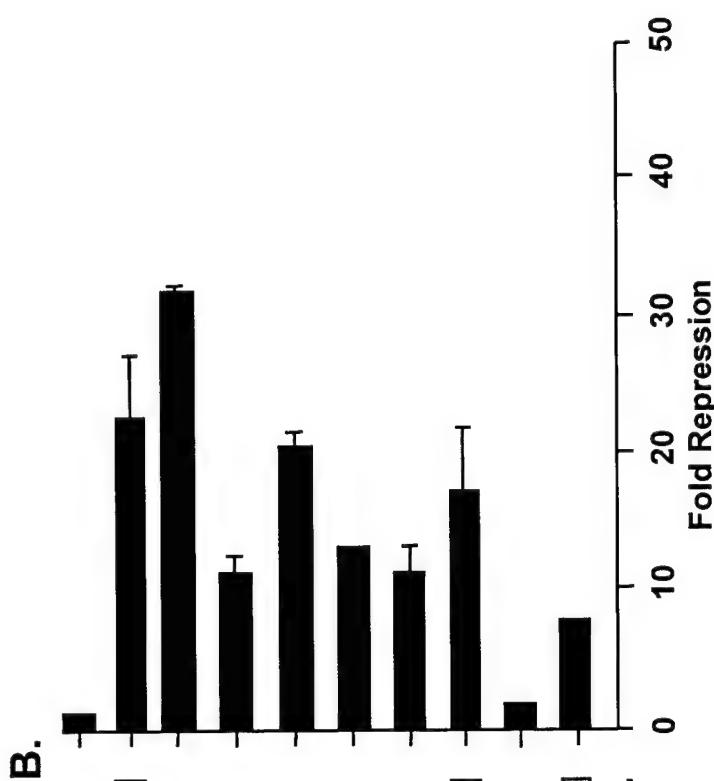
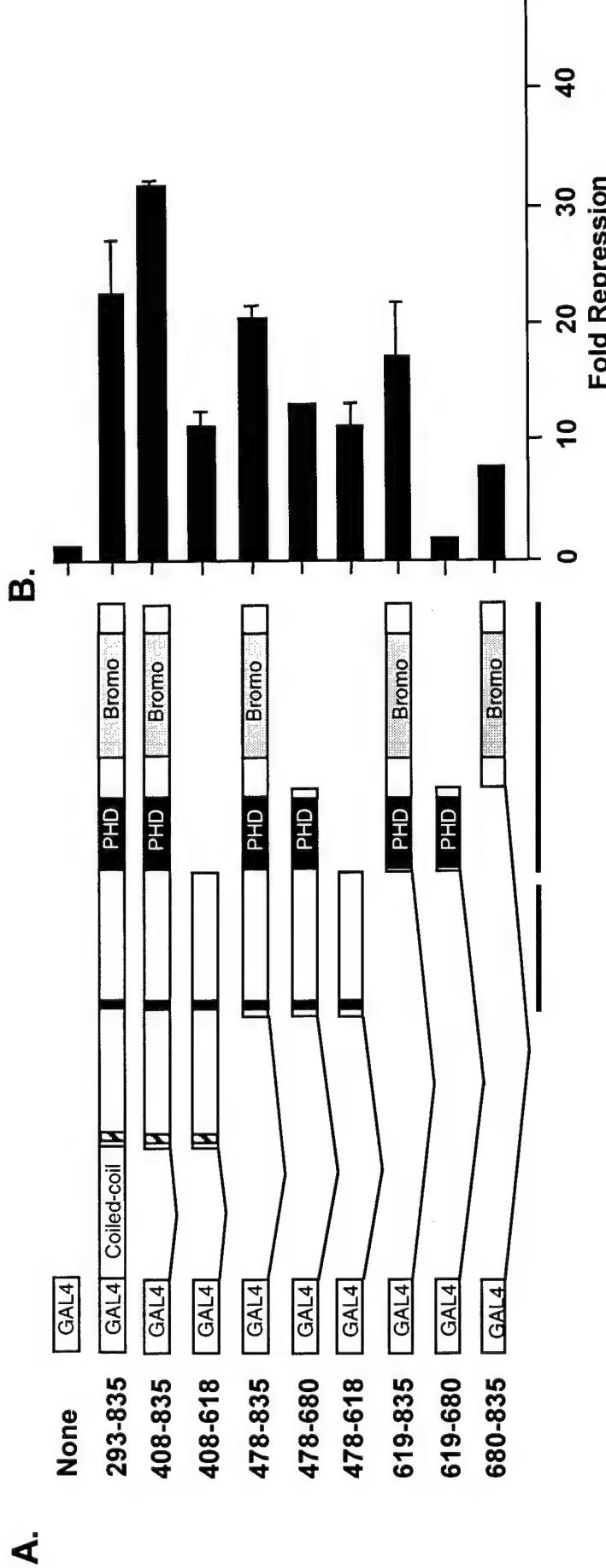
of the GAL4-KAP-1 (619-835) expression construct. (B) Amino acid sequence alignment of the KAP-1 bromodomain with related sequences from TIF1 α , TIF1 γ , WCRF180, hGCN5, and pCAF. Schematic illustration of the secondary structure and relevant position of each structural element in the bromodomain sequence is indicated above. Amino acid residues that have been conserved or where the chemical nature of the side chain has been maintained are shaded in black. Bolded amino acids in the pCAF sequence represent amino acids mutated in the study by Dhalliun et al. Asterisks mark the amino acids that were mutated in this study. The vertical line identifies the position of the 781trunc mutation. Each mutation was made in the context of the GAL4-KAP-1 (619-835) expression construct. (C) All experiments were done in NIH/3T3 cells, as described in Figure 1. Black bars represent PHD finger mutations; Grey bars represent bromodomain mutations. Stable expression of each protein was determined via transfection into COS1 cells followed by immunoprecipitation of 35 S-met labeled whole cell extracts with anti-GAL4 (DBD) antiserum (1 μ g).

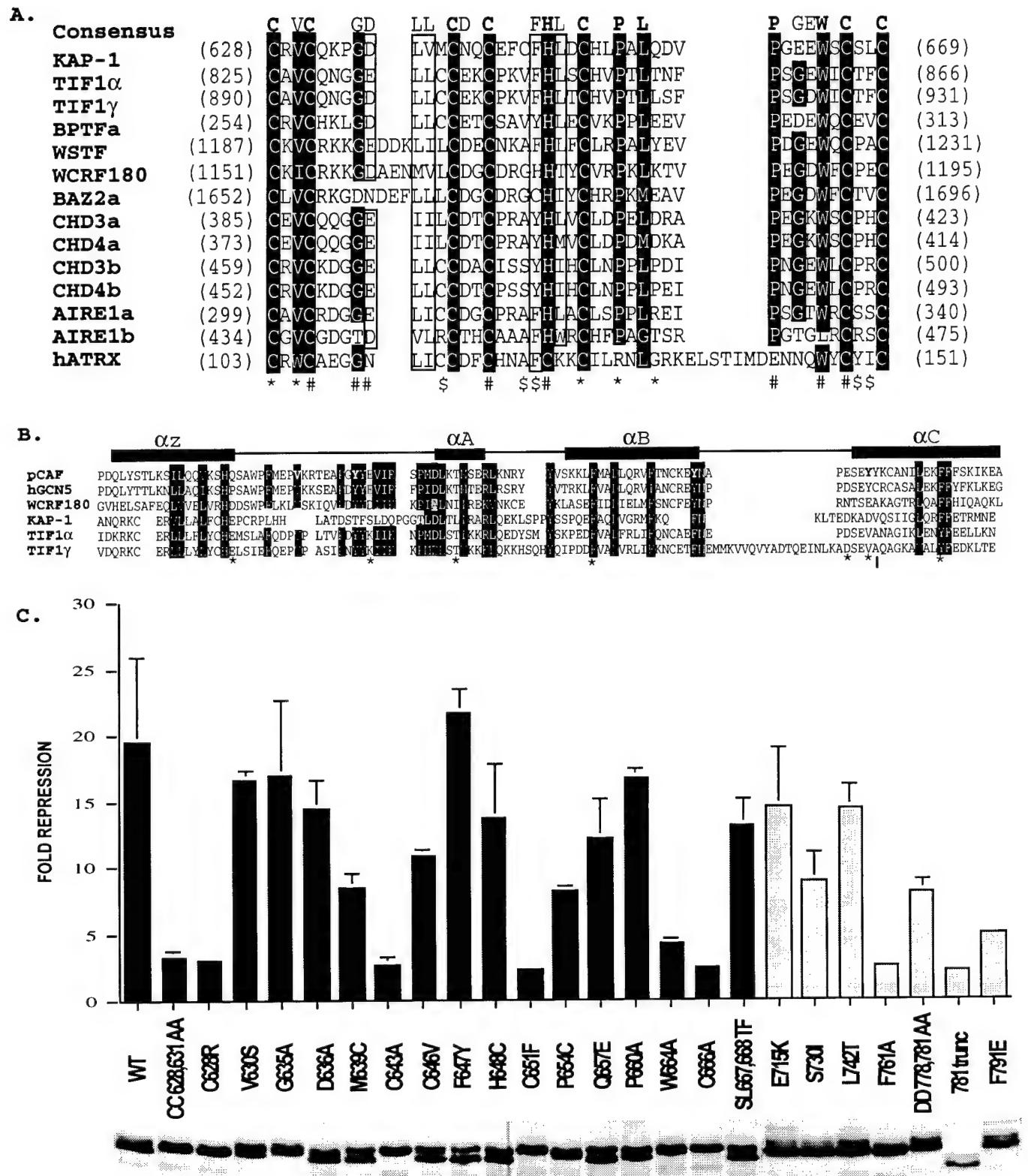
Figure 3. Heterologous chimeric fusions of related PHD fingers or bromodomains to KAP-1. (A) Schematic illustration of chimeric fusion proteins engineered. The chimeric fusion proteins possess the following amino acid sequences of the individual proteins respectively: TIF1 α /KAP-1(815-876/680-835), TIF1 γ /KAP-1(880-941/680-835), ATRX/KAP-1(210-283/680-835), Mi-2 α /KAP-1 (372-423/680-835), WCRF180/KAP-1 (1141-1205/680-835), KAP-1/TIF1 α (619-679/877-1016), KAP-1/TIF1 γ (619-679/942-1120), KAP-1/GCN5 (619-679/714-837), and KAP-1/WCRF180 (619-679/1389-1556). All experiments were done in NIH/3T3 cells, as described in Figure 1. White bars represent natural PHD finger/bromodomain fusions; Black bars represent chimeras with the KAP-1 bromodomain; Grey bars represent chimeras with the KAP-1 PHD

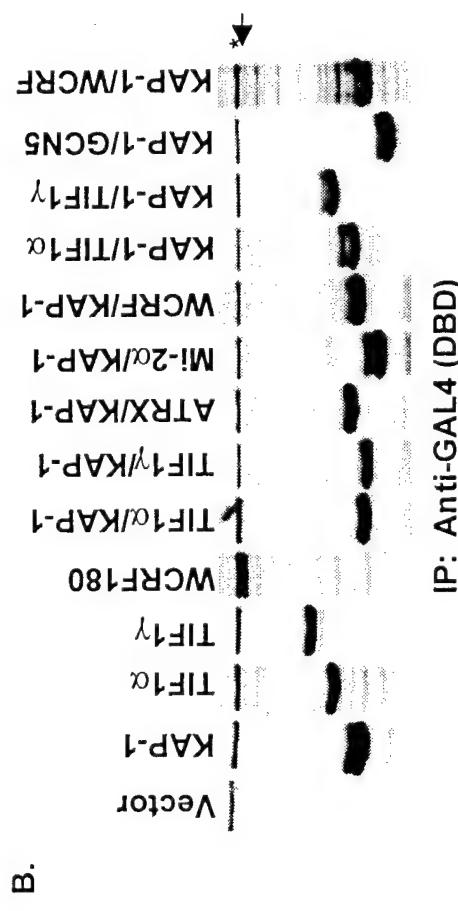
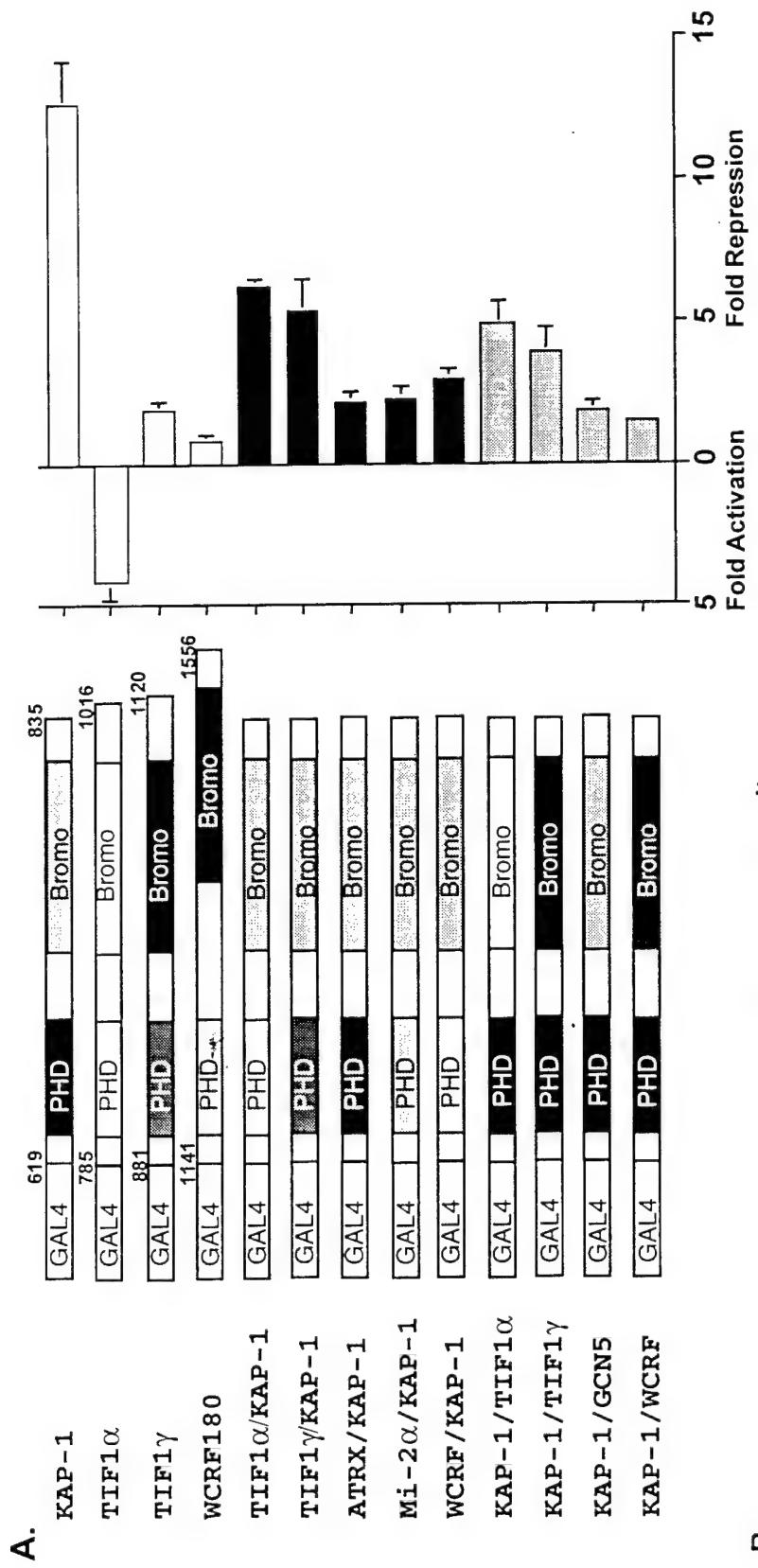
finger. (B) Each plasmid was tested for stable protein expression via transfection into COS1 cells followed by immunoprecipitation of 35 S-met labeled whole cell extracts with anti-GAL4 (DBD) antiserum (1 μ g). Arrow represents the migration of the heterologous GAL4-WCRF180 PHD finger/bromodomain. The asterisk represents a non-specific band retained during the immunoprecipitation.

Figure 4. The PHD finger and bromodomain of KAP-1 specifically interact with the COOH-terminal amino acids of Mi-2 α /CHD3. (A) The yeast strain L40 cotransformed with the indicated KAP-1 mutants as LEXA DNA binding domain fusions and the rescued Mi-2 α -ct:GAL4 activation domain clone. Positive interaction is inferred by blue yeast due to the activation of the integrated LEXA responsive *LacZ* reporter gene and subsequent hydrolysis of the synthetic substrate X-gal. (B) The interaction between KAP-1 and Mi-2 is specific for the C-terminus of the alpha isoform. Mi-2 α /CHD3 is a member of the CHD family of proteins which are characterized by the seven signature motifs of a "DEAH" box ATPase/Helicase. Other signature motifs include two PHD fingers (light grey hexagons), two chromodomains (hatched circles), and a putative DNA binding domain (stripped bar). The putative KAP-1 interaction domain (KID) is indicated. In frame fusion between the COOH-terminal sequences of hMi-2 α (accesion number 3298562), hCHD3 (accession number 2645433), hMi-2 β /CHD4 (accesion number 1107696) were designed with the GAL4 activation domain. Numbers represent the corresponding amino acid in the respective sequences. L40 yeast cotransformed with the indicated GAL4 activation domain fusion proteins and the wild type LEXA KAP-1 PHD finger and bromodomain illustrated the specificity of this interaction for Mi-2 α .

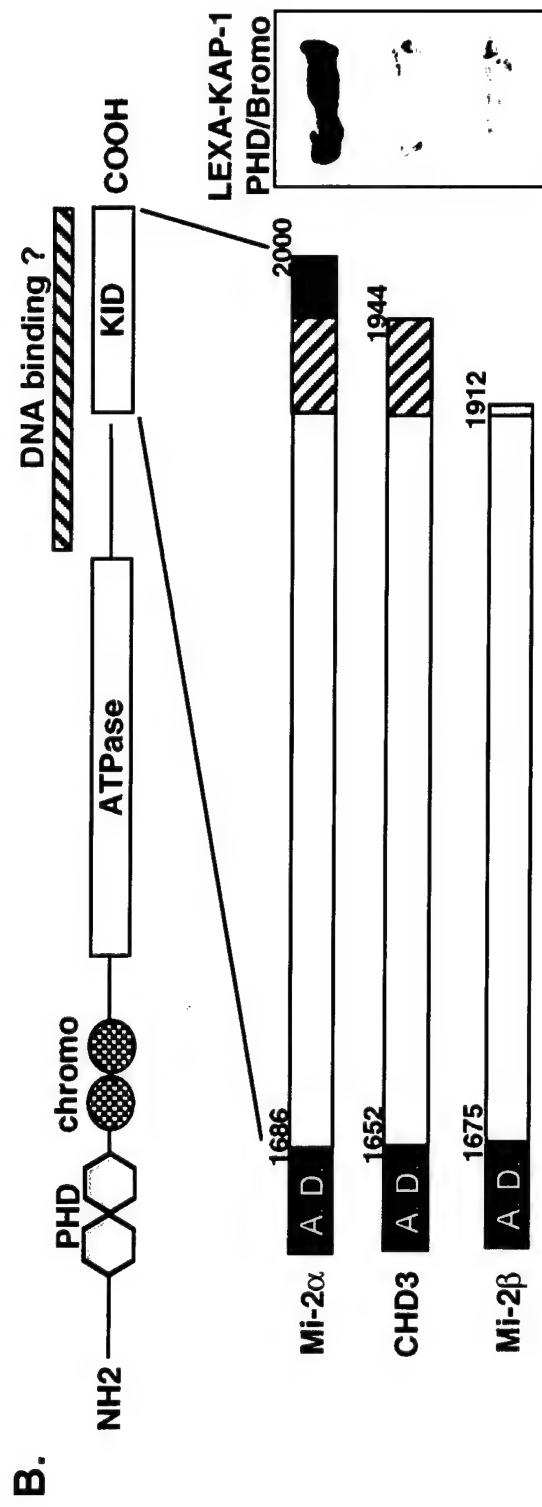
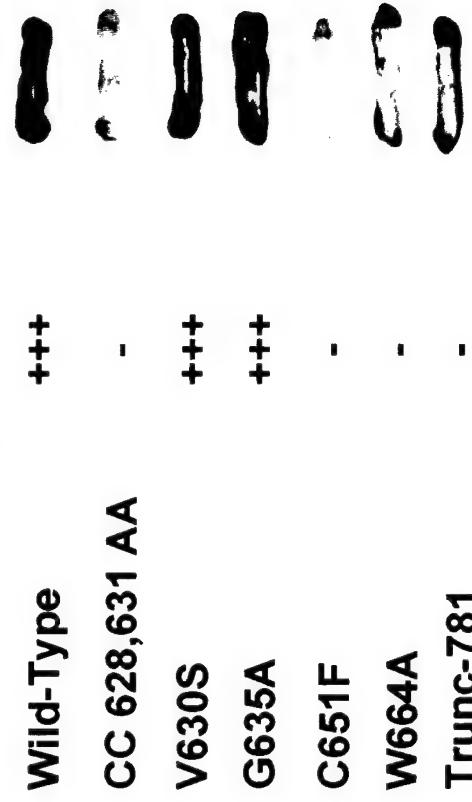
Figure 5. Role of Mi-2 α and histone deacetylase in KAP-1 repression. (A) Expression of the KAP-1 interaction domain of Mi-2 α (amino acids 1738 to 2000) dominantly inhibits KRAB-mediated repression. Schematic illustration of the expression plasmids and luciferase reporter used in transient transcription assays are displayed. Black triangles represent titrating amounts (1 μ g, 2 μ g, and 4 μ g) of plasmid DNA. Expression of protein's was confirmed in COS1 cells (data not shown). Fold repression was calculated as described in Figure 1. (B) Addition of the histone deacetylase inhibitor, TSA, partially reverses the repression activity of the KAP-1 PHD finger/bromodomain. NIH/3T3 cells were transiently co-transfected with 0.5 μ g of the indicated reporter plasmid and 5 μ g of the heterologous GAL4 expression plasmids: KAP-1 (619 to 835), SAP30, BCL6-POZ. Twenty-four hours post-transfection, the cells were treated with 300 nM TSA (Wako) for an additional 24 hours prior to harvesting. Fold repression was calculated as described in Figure 1. Fold repression in the absence of TSA (-TSA, stippled bars); fold repression in the presence of TSA (+TSA, Black bars).

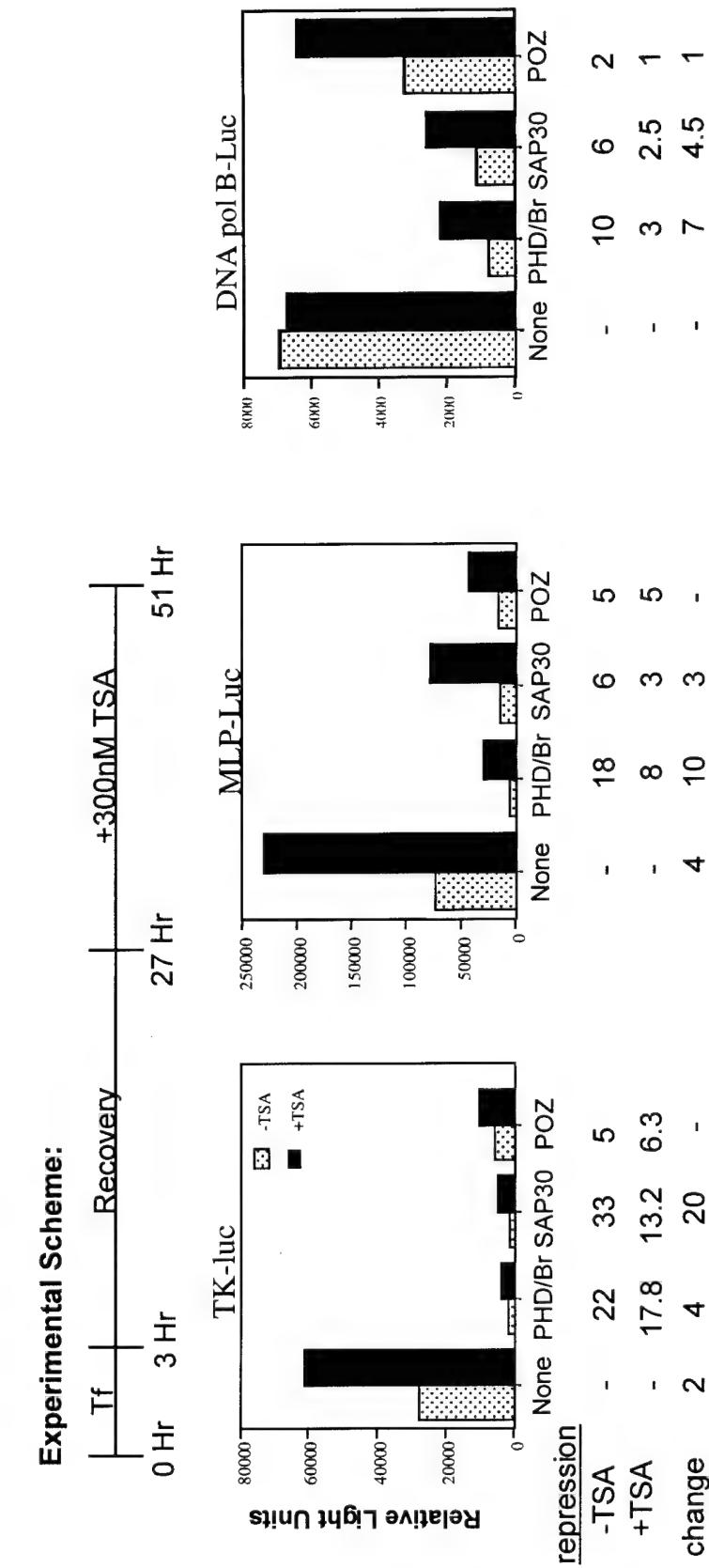
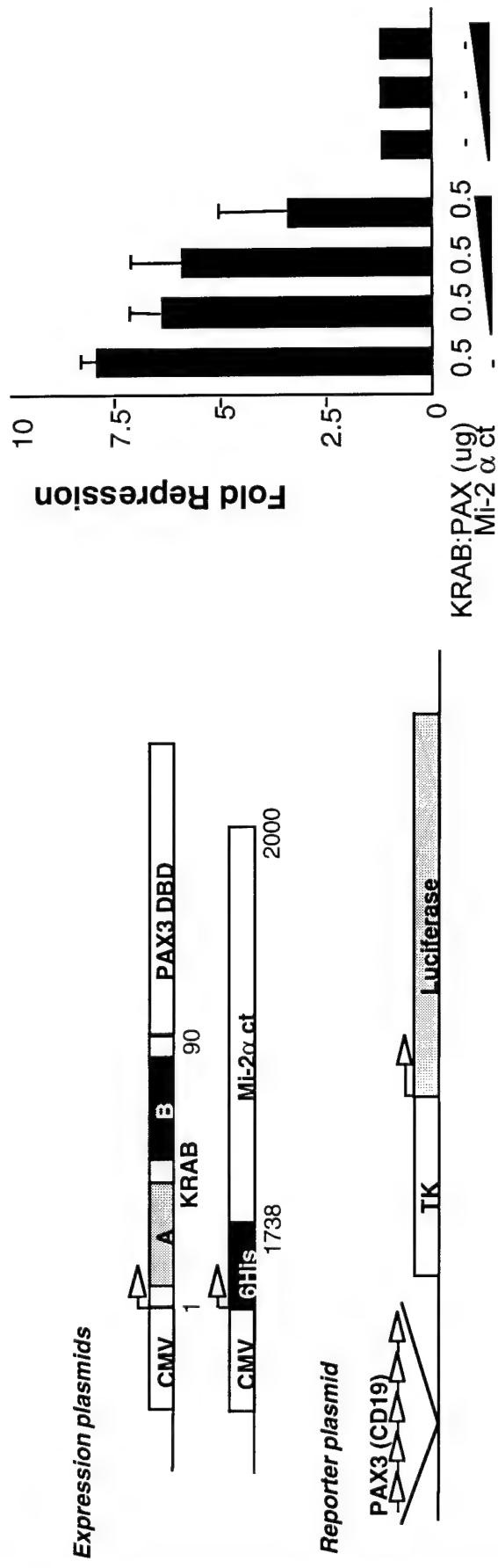




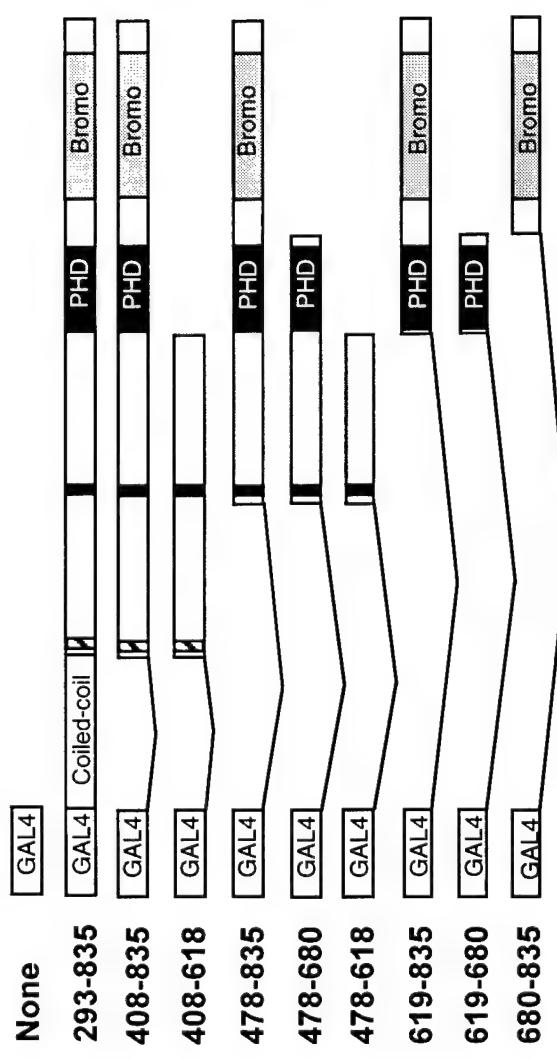


A. Repression KIP54/Mi-2 α

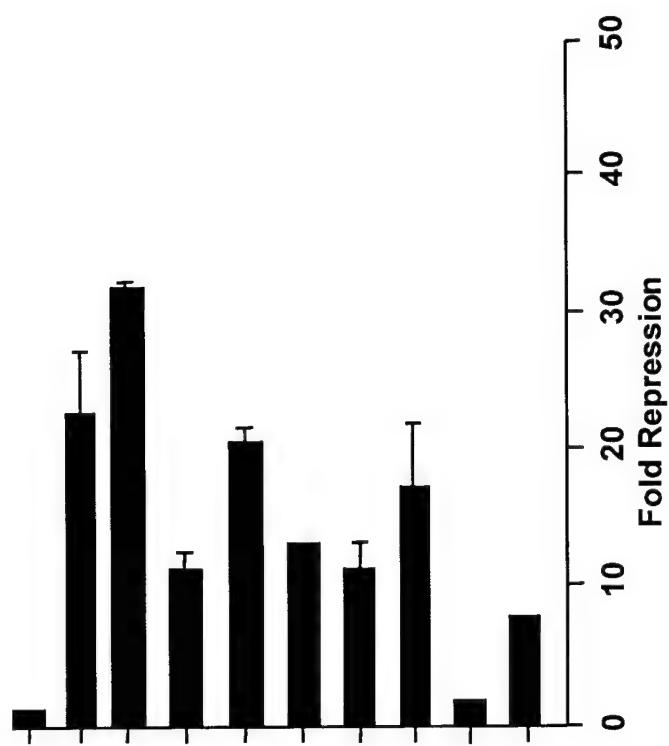


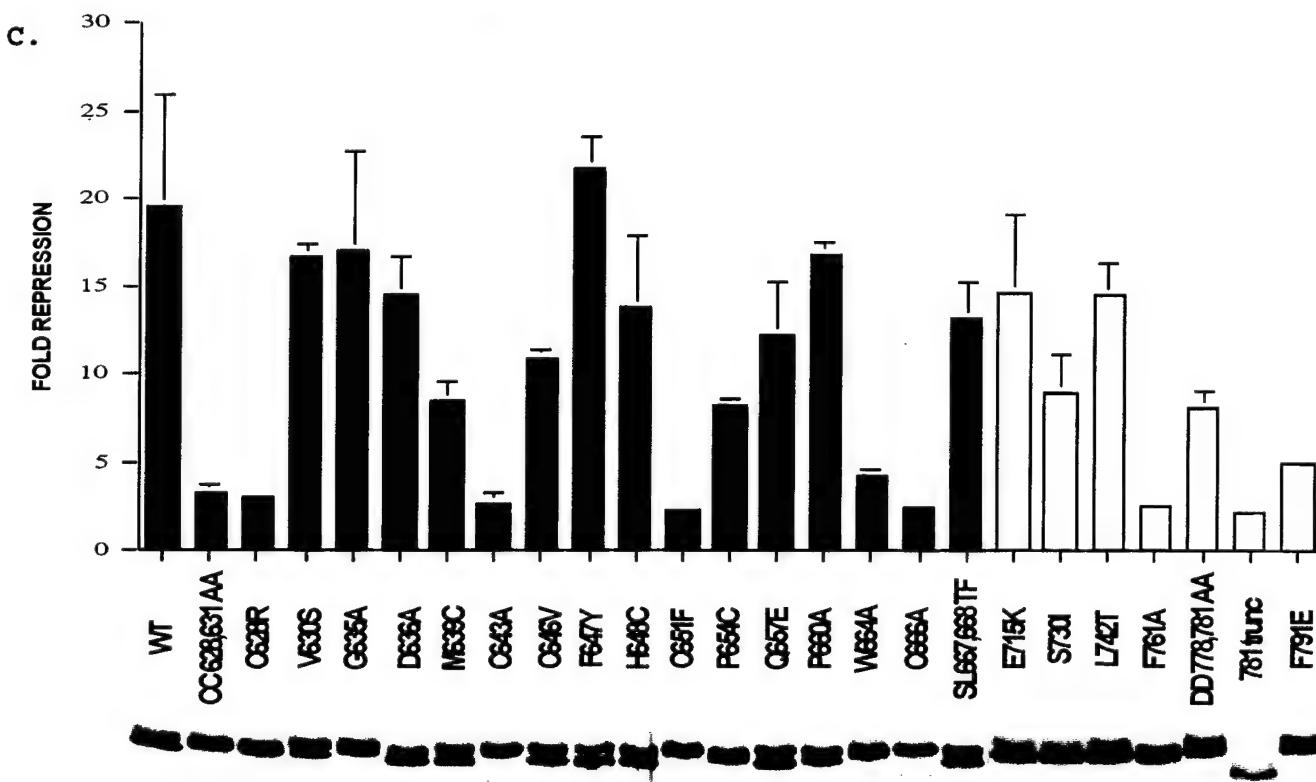
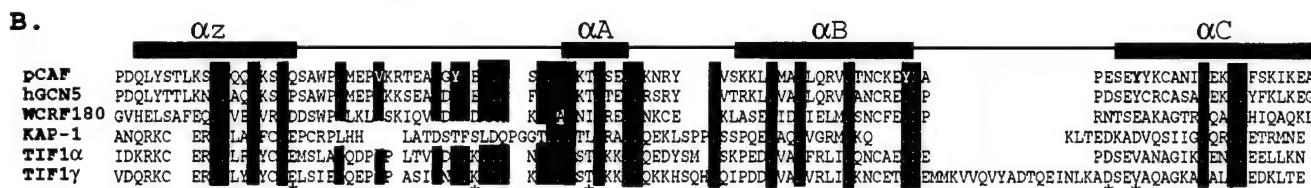
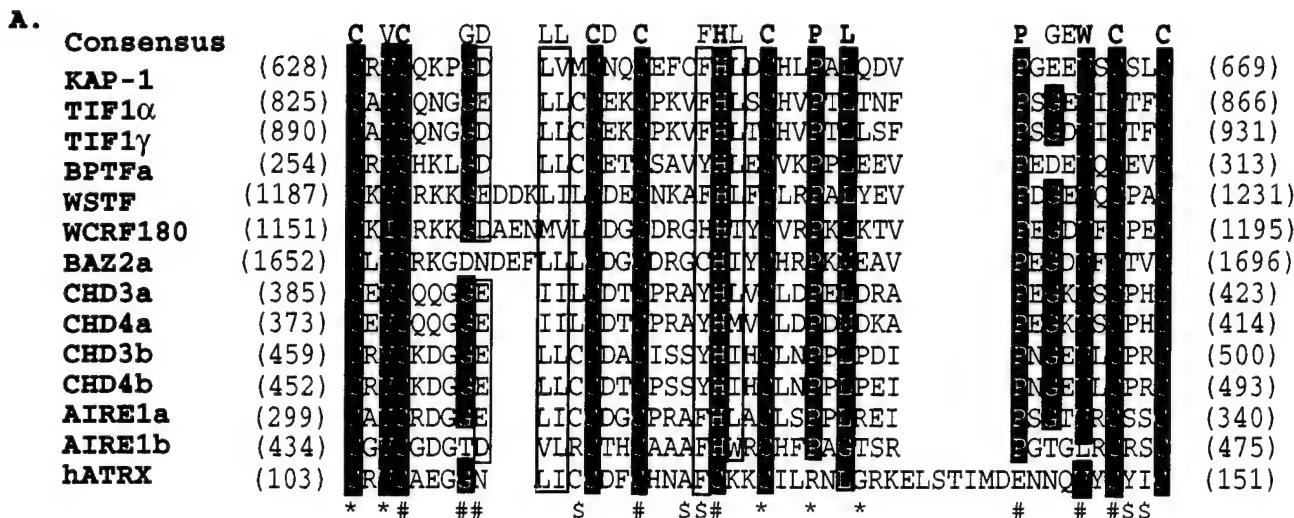


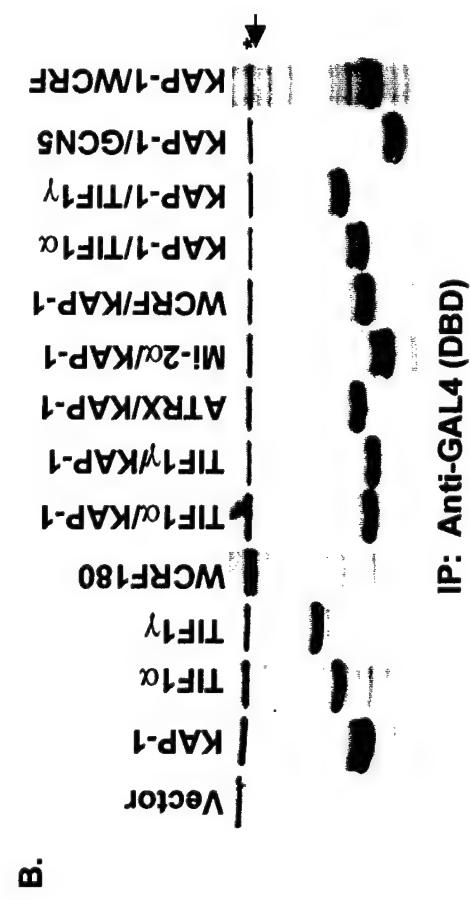
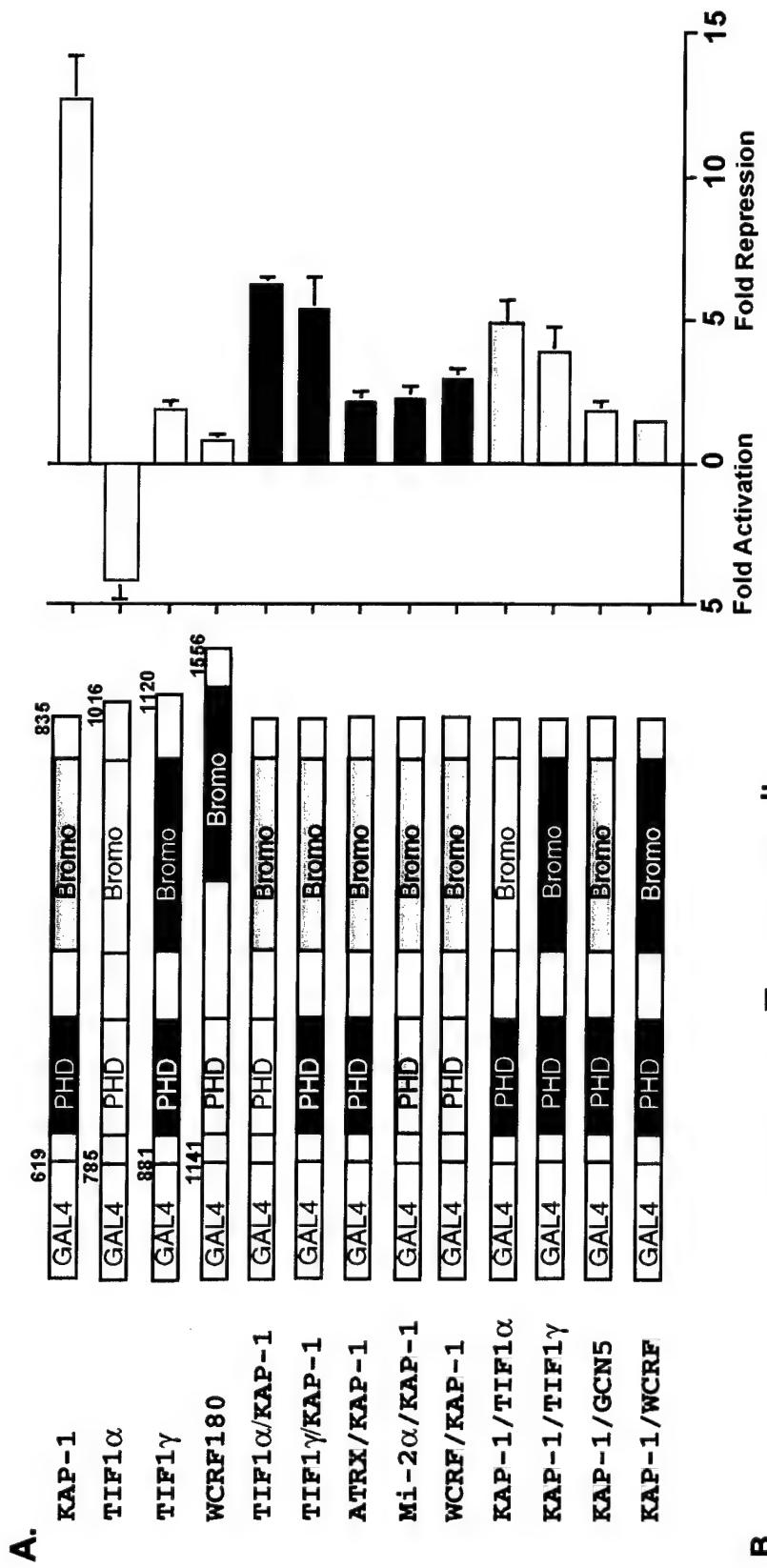
A.



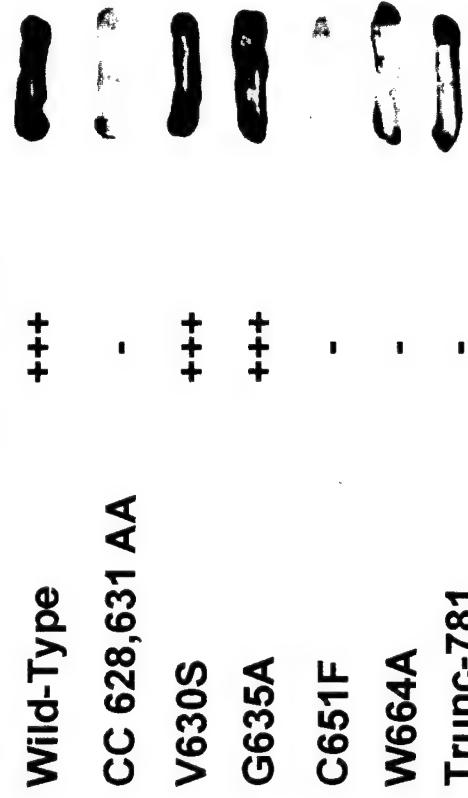
B.



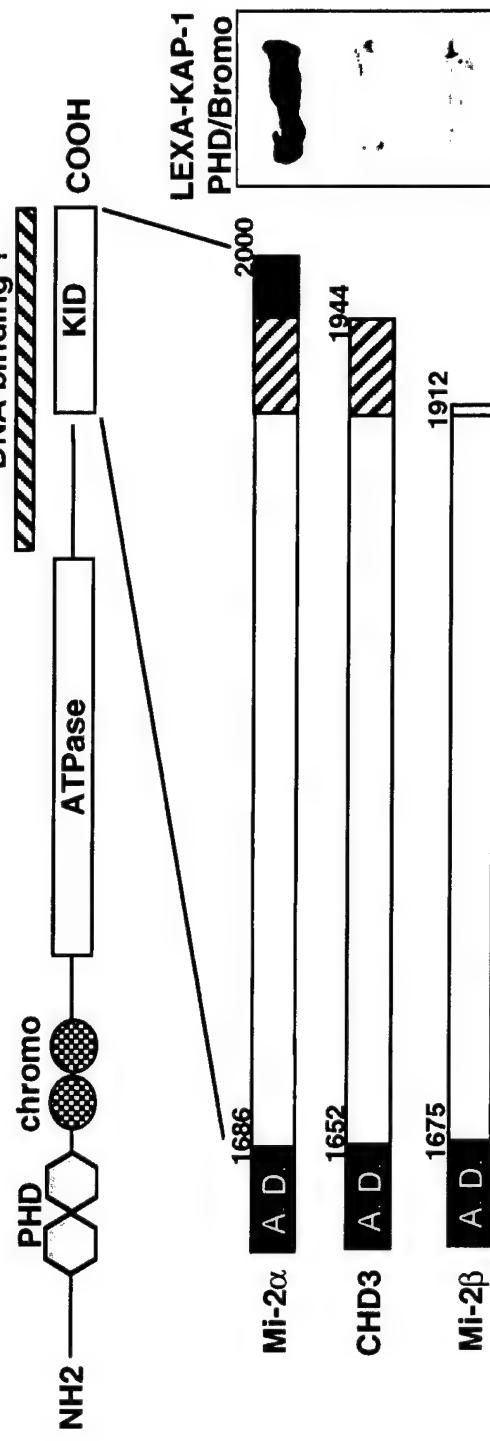


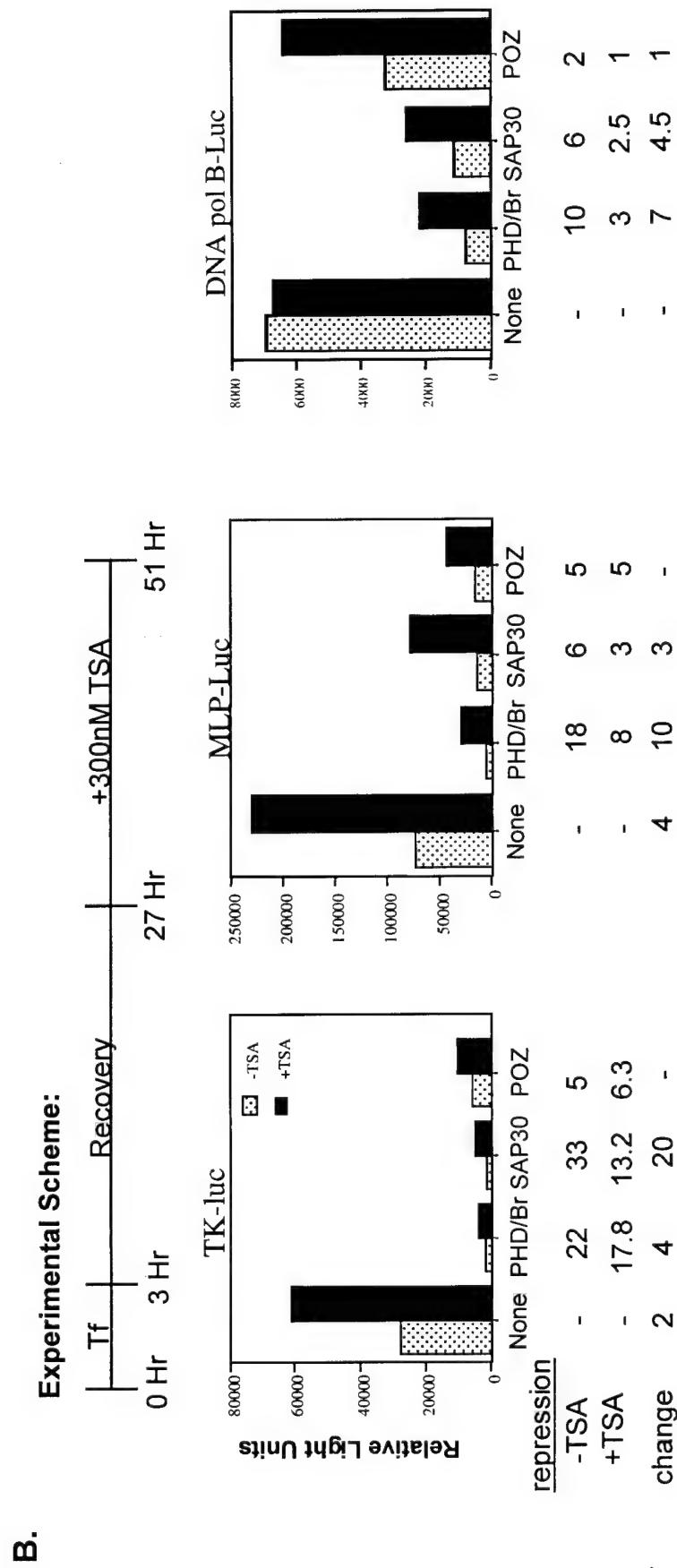
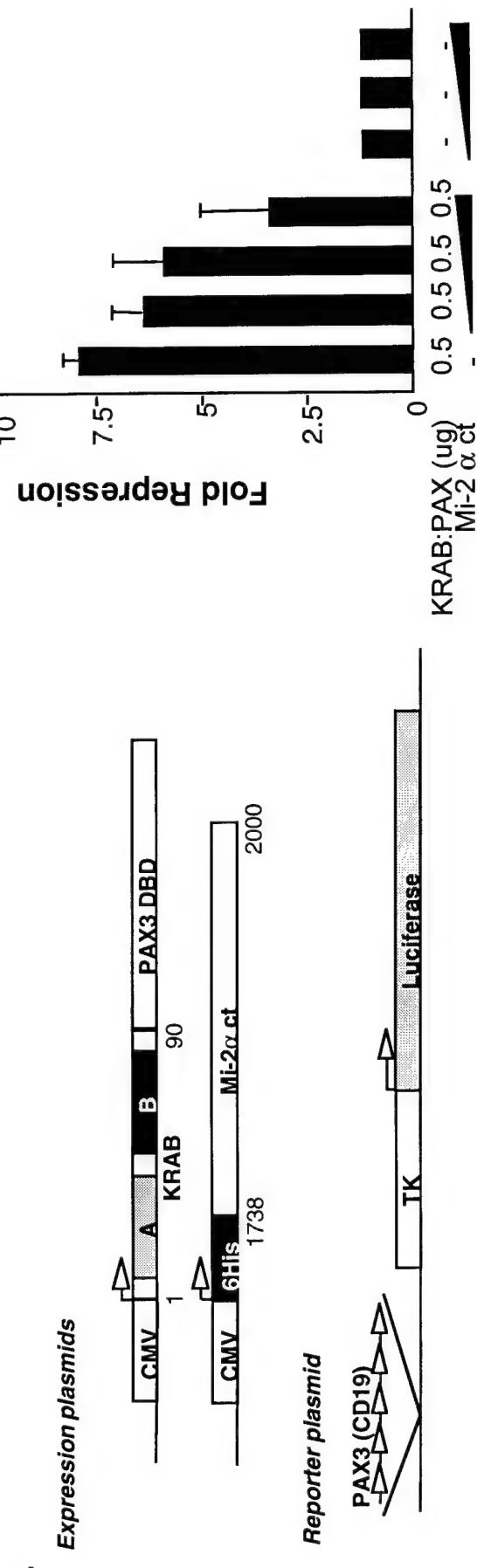


A. Repression KIP54/Mi-2 α



B. DNA binding ?





Solution structure of the PHD domain from the KAP-1 corepressor: Structural determinants for PHD, RING, and LIM zinc-binding domains.

**Allan D. Capili^{1#}, David C. Schultz^{2#}, Frank J. Rauscher III^{2*} and
Katherine L.B. Borden^{1*}**

¹Structural Biology Program, Department of Physiology and Biophysics, Mount Sinai School of Medicine, New York University, New York, NY, 10029

²The Wistar Institute, 3601 Spruce Street, Philadelphia, PA, 19104

#These authors contributed equally to this work

***Correspondence to:**

Phone: 212-659-8677

Fax: 212-864-2456

Email: kathy@physbio.mssm.edu

OR

Phone: 215-898-0995

Fax: 215-898-3929

Email: Rauscher@wistar.upenn.edu

Running title: Solution structure of the PHD domain

Character Count: 39,638

Abbreviations: NMR, nuclear magnetic resonance spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlated spectroscopy; HSQC, heteronuclear single quantum correlation; 2D, two dimensional; 3D, three dimensional; RING, really interesting new gene; PHD, Plant Homology Domain; LIM, Lin11/Isl-1/Mec-3

Abstract

PHD domains are found in over 400 eukaryotic proteins, many of which are transcriptional regulators. Naturally occurring point mutations or deletion of this domain contribute to a variety of human diseases including ATRX syndrome, myeloid leukemias and autoimmune dysfunction. Here we report the first structural characterization of a PHD domain. Our studies reveal that the PHD domain from KAP-1 corepressor binds zinc in a cross-brace topology between anti-parallel β -strands reminiscent of RING domains. By a mutational analysis, we define the structural features required for transcriptional repression by KAP-1 and explain naturally occurring, disease-causing mutations in PHD domains of other proteins. From a comparison of the PHD structure with previously reported RING and LIM structures, we infer sequence determinants which allow discrimination among PHD, RING, and LIM motifs.

(We are in the process of submitting coordinates to the PDB)

Keywords: KAP-1/LIM/PHD/RING/zinc-finger

Introduction

Biological processes depend on the spatial and temporal targeting of specific macromolecular interactions. The presence of highly conserved amino acid sequence motifs and their structural arrangement within a novel protein can provide the first clue to the protein's function. Modular motifs within multi-domain proteins often specify interacting partners, and the identification of these interactions has been beneficial in defining biochemical functions of particular proteins and their cellular pathways. Furthermore, these modular domains are often targets of mutations that disrupt the cellular pathways in which they function. Identification of modular domain type through sequence homology studies alone can lead to ambiguous domain type assignment, making prediction of function a difficult enterprise. Thus, biophysical characterization and determination of the 3D structure are essential for understanding cellular and biological function.

The plant homeodomain (PHD) or leukaemia associated protein (LAP) domain is a relatively small motif of approximately 60 amino acids that is found in over 400 eukaryotic proteins, many of which are believed to be involved in the regulation of gene expression, including the KAP-1/TIF1 β , WCRF/WSTF, Mi-2 and CBP/p300 families (Aasland *et al.*, 1995). It has been suggested that this domain is involved in protein-protein interactions related to a possible role in chromatin-mediated regulation of gene expression (Jacobson and Pillus, 1999). Although no direct role in transcriptional processes has been reported for this domain, recent reports indicate that the extended PHD domain of AF10 is necessary for homo-oligomerization, an event required for the ability of the AT hook motif in AF10 to bind DNA (Linder *et al.*, 2000).

The biological importance of the PHD domain is underscored by its involvement in the pathogenesis of several human disorders. Clinically relevant missense mutations in the PHD domain of the ATRX protein result in α -thalassemia and mental retardation (Gibbons *et al.*, 1997; Rinderle *et al.*, 1999). Germline nonsense mutations in the *AIRE* gene result in truncated proteins where one or both of the PHD domains are deleted in patients with autoimmune polyglandular syndrome type 1 (APECED) (The Finnish-German APECED Consortium, 1997). Somatically acquired point mutations in the PHD domain of ING1 have been identified in head and neck squamous cell carcinomas (Gunduz *et al.*, 2000). Chromosomal rearrangements that delete this motif in proteins such as MLL, CBP, MOZ and AF10 result in myeloid leukemias (Jacobson and Pillus, 1999). Further, genes encoding PHD domain proteins have been identified in the critical deletion regions of several contiguous gene deletion syndrome such as Williams syndrome (WSTF) and the immunodeficiency syndrome ICF (DMNT3B) (Aapola *et al.*, 2000; Lu *et al.*, 1998). The prevalence of disease-causing mutations in PHD domains suggests that they play a basic and essential role in the normal function of human cells.

The PHD domain is a motif characteristically defined by seven cysteines and a histidine that are spatially arranged in a C4HC3 consensus with intervening sequences of varying length and composition (Aasland *et al.*, 1995). This particular arrangement of amino acids is highly homologous to the RING and to the LIM domains (Figure 1A) (Borden and Freemont, 1996; Dawid *et al.*, 1994). In both the RING and the LIM domain, these conserved cysteine and histidine residues are utilized to bind two zinc atoms, a process cooperatively coupled to the folding of the domains. The LIM uses a sequential zinc ligation scheme, where the first and second pair of metal ligands creates the first zinc-binding site, while the third and fourth pair form the second site (Figure 1B). In contrast, the RING domain utilizes a unique cross-brace

ligation topology, where the first and third pair of metal ligands form the first zinc-binding site (site I), while the second and fourth pair form the second site (site II) (Figure 1B). The different arrangements of metal ligands result in dramatically different three-dimensional structures (Figure 1C). The choice of ligation topology is undoubtedly critical for proper folding and domain functionality. Given the sequence similarity amongst PHD, RING and LIM domains, one is unable to predict the ligation scheme and fold by sequence analysis alone.

Here, we report the first solution structure of the PHD domain from the KAP-1 corepressor (also known as TIF1 β or KRIP1), a universal corepressor for the KRAB-zinc finger superfamily of transcriptional repressors (Ryan *et al.*, 1999). We demonstrate that this minimal domain binds two zinc atoms in a cross-brace RING-like arrangement. Not only are the zinc ligation schemes identical, but the tertiary topologies of the PHD and RING domains are remarkably similar. From the comparison of PHD, RING and LIM domains, it is possible to derive unique structural determinants. Structure based site-directed mutational analysis reveals structural features of the PHD domain that are critical for the ability of KAP-1 to repress transcription. Furthermore, we are able to rationalize the structural consequences of mutations in the PHD domains of other proteins such as ATRX and ING1, where these mutations are correlated with human disease.

Results and Discussion

The PHD domain requires zinc for folding

The conservation of eight potential metal ligands in the PHD family suggested that it is a zinc-binding protein. In order to analyze the chemical and physical characteristics of the PHD domain, we expressed the minimal KAP-1 PHD (amino acids 618-679) in *Escherichia coli* and purified the protein to homogeneity using affinity and size exclusion chromatographies. Analytical ultracentrifugation experiments suggest that greater than 90% of the PHD protein in solution is a monomer (data not shown). To determine the identity and stoichiometry of metal bound by the PHD domain, we used inductively coupled plasma (ICP) spectrometry. These measurements indicate that a 3.0 mM solution of KAP-1 PHD protein contained 5.2 mM ($\pm 3.3\%$) zinc. This ratio of 1 protein molecule to 1.7 zinc atoms indicates that the PHD binds 2 zinc atoms per protein molecule, potentially utilizing its eight conserved cysteine and histidine residues for metal ligation.

Based on this result, we investigated the role of zinc in structure formation of the PHD domain using circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopies. The CD spectrum of KAP-1 PHD denotes a folded protein with substantial β -sheet and no significant helical content (Figure 2A). Introduction of the metal chelating agent EDTA results in complete loss of secondary structure. Addition of ZnCl₂, in excess of EDTA, results in re-folding of the domain (data not shown). Similarly, ¹H/¹⁵N-HSQC NMR experiments indicate that the KAP-1 PHD domain requires zinc for proper folding (Figures 2B and 2C). In these experiments, the PHD domain was expressed in minimal media in the presence or absence of zinc (see Materials and methods). The spectrum of the sample expressed in the absence of zinc

shows poor dispersion of amide proton chemical shifts, indicative of an unfolded protein (Figure 2B). In contrast, the sample expressed in the presence of zinc produces a spectrum revealing a substantial increase in chemical shift dispersion, a decrease in line widths, and the appearance of virtually all expected backbone amide resonances; all hallmarks of a well-folded protein (Figure 2C). Thus, like RING and LIM domains, the binding of two zinc atoms is necessary for folding of the PHD domain.

The structure of KAP-1 PHD domain reveals a cross-brace zinc ligation scheme

We determined the solution structure of the KAP-1 PHD domain using standard homonuclear and heteronuclear NMR techniques. Initial NMR studies were conducted on an N-terminal His-tagged construct of KAP-1 (amino acids 618-679) containing the entire PHD domain. A comparison of 2D $^1\text{H}/^{15}\text{N}$ HSQC and NOESY spectra of tagged and untagged proteins revealed that the two spectra were identical apart from resonances corresponding to the tag (data not shown). Thus, the presence of this N-terminal His-tag had no influence on the structure of the PHD domain.

An initial set of structures was generated without reference to metal ligation in order to determine the residues involved in zinc-binding. These structures indicated a cross-brace ligation scheme with residues C628, C631, H648 and C651 forming one zinc-binding site (site I) while C640, C643, C666 and C669 forming the other (site II). These sites were partially defined by the individual NOEs observed between residues implicated in zinc ligation. For instance, long-range NOEs were observed between C666 and C640, C631 and C651, C640 and C669, in addition, several NOEs between residues adjacent to the zinc ligands were observed. Subsequent structure calculations included Zn atoms and additional constraints defining tetrahedral

coordination. To ensure that a cross-brace was the correct zinc ligation scheme, alternative schemes were calculated but none satisfied the experimental constraints. The final ensemble is displayed in Figure 3A and all atoms depiction in Figure 3B. Statistics of the ensemble are given in Table I.

The structure of KAP-1 PHD domain reveals a globular domain binding two Zn atoms in a distinguishing cross-brace fashion (Figure 3). In this arrangement residues C628, C631, H648 and C651 form a single zinc-binding site (site I), while C640, C643, C666 and C669 form the second site (site II) (refer to Figure 3E). The formation of site I begins with an eight residue loop (residues 627-634) which continues into an extended region (residues 635-636) followed by the first β -strand (β_1 ; residues 637-640). Site II is then created by a turn (residues 641-644) that ends into the second β -strand (β_2 ; residues 645-647) running anti-parallel to β_1 . The first zinc-binding site is completed by another loop (residues 648-652). This portion is followed by an extended flexible region, (residues 653-663; shown in blue, Figure 3A) demarcated by two proline residues (P654 and P660) which appear to act as a hinge relative to the rest of the domain (Figure 3A). This flexible region is followed by a third β -strand (β_3 ; residues 664-666) leading to the completion of the second zinc-binding site (residues 666-670). Numerous hydrophobic interactions appear to stabilize this structure. F647 is at the centre of the hydrophobic core packing against W664, V638, L656, and H652. Interestingly, the proton resonances for the aromatic ring of F647 are not degenerate indicating that the phenyl group is not free to rotate in the core suggesting that the core is tightly packed. Analysis of the electrostatic surface potential reveals two areas of discrete charge (Nicholls *et al.*, 1991). Most notably, the flexible hinge (residues 653-663) is lined with several Glu and Asp residues contributing to a negatively charged surface and the presence of Arg629 and Lys633 results in a positively charged cluster

around site I. Ongoing mutational analyses will be utilized to determine the importance of these features to KAP-1 PHD function. In terms of charge and topology, no structural features typical of nucleic acid binding proteins are observed, such as those present in the Kruppel-like or GATA-like zinc fingers (Klevit *et al.*, 1991; Omichinski *et al.*, 1993). This suggests that the PHD is not a nucleic acid binding domain.

The PHD domain structurally resembles a RING

Unlike the sequential zinc ligation pattern used by the LIM domains, the KAP-1 PHD domain structure shows a zinc ligation pattern similar to RING domains. Therefore, we carried out a detailed comparison between the PHD domain and two previously reported RING domain structures, from the promyelocytic leukemia protein PML and the immediate early equine herpes virus protein IEEHV (Protein Data Bank codes 1bor and 1chc, respectively; (Barlow *et al.*, 1994; Borden *et al.*, 1995a)). These proteins have no sequence homology outside of conserved metal binding residues. Strikingly, in all three cases the inter-zinc distance is ~14Å, presumably because the central β -strand of these molecules is the same length (β_2 in KAP-1 PHD) and all use the cross-brace ligation scheme. Because PHD and RING domains use different permutations of cysteines and histidines for metal ligation, for clarity we denote the i^{th} conserved metal-ligating residue along the primary sequence by, ml_i . Notably, a conserved hydrophobic core residue (F647) is located in β_2 , just N-terminal to ml_5 . Inspection of these structures indicates that the first zinc-binding site and anti-parallel β -strands ($\beta_1\beta_2$) of PHD overlay strikingly well with the two RING structures (Figure 4). An α -carbon backbone alignment of RING and KAP-1 PHD structures for site I overlay nearly exactly, as do the zinc atoms. The root mean square difference (RMSD) of the α -carbon backbone for residues around site I is 1.4Å when compared

to PML and 1.3Å when compared to IEEHV (Figure 4B). Both PML and IEEHV RINGs have a large insertion along the β_1 strand relative to KAP-1 PHD (Figure 5). Comparison of residues just N-terminal to ml_1 through to just C-terminal to ml_6 , which includes residues in the hydrophobic core, results in RMSDs between KAP-1 and PML of 2.0Å and between KAP-1 and IEEHV of 2.1Å. The sequence identity in this region is 20% between PML and KAP-1 and 22% between IEEHV and KAP-1. Normally, RMSD values of ~2Å are expected for proteins with sequence identities greater than 30% (Flores *et al.*, 1993). Thus, these levels of similarity are striking considering there is virtually no sequence conservation between these domains outside of their conserved zinc ligands.

Like the PHD family, the RING family shows substantial structural plasticity between ligands ml_6 and ml_7 where the chain completes site I and continues on towards site II. In PHD domains, the spacing varies from 12 to 46 residues, and in RING domains, the spacing varies from 4 to 48 residues. The RINGs of IEEHV, RAG1, and BRCA1 have large inserts in this region, forming a helix. However, RINGs with shorter inserts, e.g. PML and BARD1, do not have a helix in this region (information for BRCA1 and BARD1, personal communication, P. Brozvic and R. Klevit; reviewed in Borden, 1998). In contrast, this region in the KAP-1 PHD contains the flexible hinge (residues 653-663). Although this insert is comparable in size to that of RINGs containing a helix, the propensity for this region to adopt a helical conformation is highly unlikely due to the presence of two proline residues (P654 and P660).

A major structural difference between PHD and RING domains is found in the hydrophobic core. Residues required for the hydrophobic core of the RING structures are N-terminal to ml_5 and C-terminal to ml_6 (L76 and L81 in PML, F28 and I33 in IEEHV, see also Figures 4C and 5) (Barlow *et al.*, 1994; Borden, 1998; Borden *et al.*, 1995a). Conserved

hydrophobic residues at these exact positions are also found in the PHD (F647, N-terminal to ml_5 ; and H652, C-terminal to ml_6). However, there exists an additional conserved hydrophobic residue in PHD, a tryptophan two residues N-terminal to ml_7 (W664 in KAP-1; Figures 4C and 5). This tryptophan, which is highly conserved throughout the PHD family, inserts between the core residues N-terminal to ml_5 and C-terminal to ml_6 . In RINGs, no hydrophobic core residue exists at the position two residues before ml_7 (Figures 4C and 5). Rather, in all RING structures reported, the amino acid at this position is actually solvent exposed (reviewed in Borden, 1998). This additional core residue appears specific to the PHD family. Because of the presence of W664, the side-chains of F647 and H652 are oriented differently relative to the RING, although the trace of the main-chain is virtually unaltered relative to the RING fold (Figure 4C). Ligands ml_3 and ml_4 in site II are positioned nearly identically in both PHDs and RINGs.

Aside from differences in the hydrophobic core, there are significant differences in the main-chain trace between ml_6 and ml_7 in PHD compared to RING domains. At H652, just C-terminal to ml_6 , the main-chain trace significantly deviates from RING domains. This difference appears to result from the participation of W664 (two residues N-terminal to ml_7) in the hydrophobic core. This alters the position of the ml_7 and ml_8 relative to the RINGs. These differences suggest that the PHD domain, although structurally similar to the RING in part, could bind structurally distinct partners and display unique biological activities. We expect that insertion size and sequence variability between ml_2 to ml_3 and ml_6 to ml_7 results in substantial structural plasticity in PHD domains, imparting specificity onto a common structural scaffold stabilized by zinc.

Sequence Determinants for RING, LIM and PHD domains

The structural similarity between PHD and RING domains, coupled to the dissimilarity with LIM domains, prompted us to examine whether there were any structure-based sequence-determinants which would enable one to predict if a given amino acid sequence is PHD, RING or LIM-like in structure. The major structural determinant appears to be based on the zinc ligation scheme utilized. Sequence alignments for selected members of the PHD, RING, and LIM families are given in Figure 5. For RING and LIM domains, conserved hydrophobic core residues determined through inspection of the structures are shown. In these cases, residues which form the hydrophobic core are shown in yellow, site I in magenta, and site II in blue. These alignments suggest that proteins with conserved hydrophobic residues at positions N-terminal to ml_5 and C-terminal to ml_6 utilize a cross-brace ligation scheme as in the PHD/RING structures whereas proteins with hydrophobic residues N-terminal to ml_3 and C-terminal to ml_4 utilize a sequential LIM-like ligation scheme. Alignments indicate that of the PHD/RING cross-brace schemes, the presence of a tryptophan residue two residues N-terminal to ml_7 may be diagnostic of a PHD fold. In contrast, LIMs have additional conserved hydrophobic core residues which are not found in PHD or RING (Figure 5). Furthermore, residue spacing in the LIM domain is much more highly conserved than for PHDs or RINGs (Figure 1A).

For PHD, RING, and LIM, flexibility in ligating residue type has been observed, e.g. His can replace Cys and even Asp can be used as a ligand. Thus the position of specific types of zinc ligands is unlikely to be a primary determinant of the zinc ligation pattern and therefore the fold of the molecule. To test this hypothesis, we swapped the type of zinc ligand, e.g. His for Cys, at ligand position ml_5 , and determined the ability of KAP-1 to act as a transcriptional repressor (Figure 6B, Table II). This mutation has little effect indicating that the type of ligating residue, at least at this position, is not crucial. However, substitution of Ala for His at ml_5 completely

destroys the transcriptional repression properties of the KAP-1 PHD domain (Figure 6B, Table II).

Structural integrity of the PHD domain is essential for KAP-1 to repress transcription

In order to define structural features of the PHD domain important for the transcriptional repression by KAP-1, mutational analysis was carried out on a GAL4-KAP-1 (619-835) expression construct (Figure 6A). This construct encodes both the PHD domain and the Bromodomain of KAP-1 which together comprise an independent repression domain (Figure 6 panels B and C, Table II, and Schultz et al., submitted). Mutations of zinc ligands in either site I (CC628, 631AA or H648A) or site II (C643A and C666A) abolished transcriptional repression (Figure 6, Table II), while mutation of the non-zinc ligating cysteine (C646V) shows no effect. A mutation which potentially introduces an extra zinc ligand (M639C) results in a lowered activity, presumably because this ‘new’ ligand may disrupt the formation of the cross-brace by competition for zinc. Further, mutations that disrupt the hydrophobic core of the molecule, W664A, are as debilitating as mutations that disrupt zinc ligation. Consistently, conservative non-disruptive mutations in the core, F647Y, do not alter transcriptional repression by KAP-1. Thus, the ability of the KAP-1 PHD domain to bind zinc and maintain the hydrophobic core are both required for its proper folding and ability to repress transcription.

The potential structural consequences of PHD domain mutations in other proteins

Naturally occurring mutations and deletions in the PHD domain of several proteins exemplify the biological importance of this domain. Germline missense mutations in the PHD domain of the ATRX protein predispose individuals to α -thalassemia and mental retardation

(Gibbons *et al.*, 1997). Somatic missense mutations found in the PHD domain of ING1 are detected in several types of carcinomas (Gunduz *et al.*, 2000). Analysis of these mutations in light of the KAP-1 PHD structure reveal the basis for their devastating nature (Table II). In the ING1 PHD domain, two mutations of site I are found in squamous cell carcinomas (Gunduz *et al.*, 2000). C631S (where this amino acids are numbered in the context of the KAP-1 PHD), like previous zinc-ligand mutations, should destroy zinc-binding and structure formation. N632S, adjacent to ml_2 , may indirectly interfere with zinc-binding and therefore, formation of site I. Two mutations in ATRX (C628R and C651F) disrupt the ability of this domain to ligate zinc. These mutations are expected to result in an unfolded protein, and equivalent mutations in KAP-1 PHD domain results in a loss of repression (Table II). Another mutation in ATRX introduces an additional potential metal-ligand (P654C) which may disrupt biological activity through misfolding. Some surface mutations (V630S and Q657E) have no effect on transcriptional activity in these assays; however, these mutations are found in ATRX patients. Thus, these residues may be required by the ATRX protein for mediating protein-protein interactions. An interesting feature of the wild-type ATRX PHD domain is that it is a PHD domain with eight cysteines (His at ml_5 is replaced by Cys). Mutation of this ligand in KAP-1 (H648C) resulted in only a slight loss of transcriptional activity (Figure 6B) suggesting that either a histidine or a cysteine can be tolerated as a metal-ligand at this position.

RING-like biological activities of the PHD domain

Since KAP-1 PHD has many "RING-like" features including the cross-brace zinc ligation scheme and a high level of structural similarity to the anti-parallel β -strands and zinc-binding site I, we conducted studies to determine whether this domain had any RING-like biological activity.

Features characteristic of RINGs include: (I) requirement of zinc for folding and biological activity and (II) formation of large multi-protein complexes by acting as scaffolds for protein-protein interactions (Borden, 2000; Kentsis and Borden, 2000). Like RING, we show that KAP-1 PHD requires zinc for folding and function. KAP-1 PHD participates in protein-protein interactions with Mi-2 α (Schultz et al., submitted), a component of the NuRD histone deacetylase complex (Brehm et al., 1999). Therefore, KAP-1 PHD may be important for proper spatial and temporal scaffolding of the repressor complex. This association is likely to be critical for transcriptional repression by KAP-1. Consistent with our results, recent work indicates that the AIRE PHD domain is required for formation of nuclear protein complexes (Rinderle et al., 1999). Further, the AF10 PHD domain is reported to be a homo-oligomerization module (Linder et al., 2000). Like RINGs, the PHD domain appears to function as a protein-protein interaction motif. Presumably, disruption of the structural integrity of PHD domain by mutation of critical residues fundamentally disrupts this action.

Recent studies have shown that some RING domains function as E3 ubiquitin ligases (reviewed in (Borden, 2000)). E3 ubiquitin ligases are components of the ubiquitin degradation pathway. Recent structural data indicate that a helix between ml_6 and ml_7 is involved in making the major contacts between the Cbl RING and the ubiquitin machinery (Zheng et al., 2000). Because of the similarity to the RING, we tested KAP-1 PHD in this system. However, the PHD domain has no detectable E3 ligase activity (data not shown). This result is consistent with our structural data which indicate that there is no helix present between ml_6 and ml_7 . It will be necessary to test several other PHD domains before this activity can be ruled out for the family as a whole.

Conclusions

We present the first structure of a PHD domain and show it is an autonomously structured domain that requires zinc for folding. From sequence analysis alone, it is impossible to determine how the eight conserved metal ligands are utilized to bind zinc. This uncertainty is evident from a recently modeled structure of the DNMT3L PHD domain where the ligation scheme was assumed to be LIM-like (Aapola *et al.*, 2000). The actual ligation of zinc by the PHD domain is accomplished by a cross-brace motif, like that used by the RING family. Not only are the ligation schemes similar between the two domains, but the structures are topologically identical from ml_1 to ml_6 , which includes the first zinc-binding site and the anti-parallel β -strands. We identify structural determinants that distinguish ligation schemes of RING, LIM and PHD domains. In the structure of the KAP-1 PHD domain, there exists a flexible hinge region between ml_6 and ml_7 characterized by a discrete negatively charged surface. Amongst the RING domains, this same region shows structural variability, forming an extended strand or a helix. The plasticity of this region may underlie the biological specificity of these rather similar domains. The PHD domain is required for transcriptional repression by KAP-1 and we present structural determinants required for this activity, which elucidate the structural basis of disease-causing missense mutations in ATRX and ING1.

Materials and methods

Preparation of PHD domain

A DNA fragment encoding the PHD domain of human KAP-1 (amino acids 618-679) was subcloned into the pQE30 expression vector (Qiagen), and expressed in *E. coli* BL21(DE3) cells

(Novagen). The His-tagged protein contained an additional 17 N-terminal amino acids, MRGSHHHHHGSDIIDE, amino acids 618-679 of hKAP-1, and 9 C-terminal amino acids, VDLQACKLN. For the untagged KAP-1 PHD domain protein, a DNA fragment encoding amino acids 618-679 was subcloned into the expression vector pQE50 (Qiagen), and expressed in *E. coli* SG13009 cells (Qiagen). The expressed protein contained 4 NH₂-terminal amino acids, MRGS, followed by amino acids 618 to 679 of hKAP-1. His-tagged or untagged protein was prepared from logarithmically growing bacteria cultured in either 2YT media or in minimal media, containing ¹⁵NH₄Cl as the sole nitrogen source and supplemented with 100 μM zinc acetate (for uniformly labeled ¹⁵N-protein). In both types of media, protein expression was induced with 1 mM IPTG for 3 hours at 37°C. Induced bacterial pellets were lysed in 20 mM NaH₂PO₄, pH 7.5, 500 mM NaCl, 5mM DTT. The His-tag PHD domain was purified by affinity chromatography on a Ni²⁺-NTA-agarose column (Qiagen), followed by fractionation through a Superdex 75 size exclusion column (Pharmacia). The untagged PHD domain was purified sequentially by ion exchange chromatography (DEAE-cellulose), hydrophobic interaction chromatography (Butyl-sepharose), and size exclusion chromatography (Superdex 75). NMR experiments utilized both constructs, while metal binding analysis (i.e. ICP and CD) used only the untagged PHD construct.

Site-directed mutagenesis

Site-directed point mutations in the KAP-1 PHD domain were engineered by standard overlap extension PCR-mediated mutagenesis procedures. The mutagenic primers for the described mutations (Figure 6) contained the following codons: CC628,631AA, TGC to GC; C628R, TGC to CGC; V630S, GTC to AGC; G635A, GGC to GCC; D636A, GAT to GCT; M639C, ATG to TGC; C643A, TGT to GCT; C646V, TGT to GTT; F647Y, TTC to TAC; H648C, CAC to TGC;

C651F, TGT to TTT; P654C, CCG to TGC; Q657E, CAG to GAG; P660A, CCA to GCA; W664A, TGG to GCG; C666A, TGC to GCC; SL667,668TF, TCA and CTC to ACA and TTC. Appropriate reading frame fusions and integrity of flanking sequences for all constructs created by PCR was confirmed by DNA sequence analysis of both strands.

Transient transfections

Protein expression from all plasmids was confirmed by transient transfection of COS-1 cells followed by immunoprecipitation of 35 S-methionine labeled cell extracts. All transcription assay transfections were done as previously described (Ryan *et al.*, 1999).

Inductively coupled plasma spectrometry

Inductively coupled plasma spectrometry was conducted using the untagged PHD sample at Galbraith Laboratories, Inc. (Knoxville, TN, USA). 25 μ l of a 3.0 mM PHD sample (20 mM NaH₂PO₄, 500 mM NaCl, 5mM DTT, pH 7.5) was prepared for ICP analysis using a wet ash digestion procedure previously described (Bock, 1979). Measurements were made on a Perkin-Elmer P2000 using a primary wavelength of 213.856 nm (Wallace, 1981).

Circular dichroism spectroscopy

Spectra were collected on a Jasco 810 spectropolarimeter at 25° using a 1 mm path-length cuvette. The concentration of the untagged PHD sample was 30 μ M in 20 mM NaH₂PO₄ and 50 mM NaCl, pH 7.5. After an initial spectrum was collected, a 6-fold molar excess of EDTA was added and pH was adjusted to pH 7.5. The sample was then left at room temperature for 16 hours to ensure it had reached equilibrium. Spectra were then recorded.

NMR spectroscopy

For NMR, KAP-1 PHD domain was 1.5-3.0 mM in 20 mM NaH₂PO₄, 500 mM NaCl, 5 mM DTT at pH7.5. No spectral changes were observed in this sample concentration range. Spectra

were recorded at 30°C on a Bruker DRX500 spectrometer. Sequential assignments were obtained using 3D $^1\text{H}/^{15}\text{N}$ NOESY-HSQC, 3D $^1\text{H}/^{15}\text{N}$ HMQC-NOESY-HSQC, and 3D $^1\text{H}/^{15}\text{N}$ TOCSY-HSQC spectra. Side-chain resonances were assigned using 3D $^1\text{H}/^{15}\text{N}$ TOCSY-HSQC, 2D ^1H -TOCSY, and 2D ^1H DQF-COSY spectra. Aromatic protons were assigned through 2D ^1H -TOCSY, 2D ^1H -NOESY, and 2D ^1H DQF-COSY spectra recorded in $^2\text{H}_2\text{O}$. 2D data were collected and analyzed as described (Borden *et al.*, 1995a; Borden *et al.*, 1995b). For side-chain and aromatic assignments, a construct with no His tag was used. NOE-derived distance restraints were obtained from 3D $^1\text{H}/^{15}\text{N}$ NOESY-HSQC spectra (τ_m of 75 and 150 ms) and 2D ^1H -NOESY spectra (τ_m of 75 and 150 ms). ϕ -angle restraints were determined based on the $^3J_{\text{HN},\text{H}\alpha}$ coupling constants measured in a 3D HNHA spectrum. Slowly exchanging amide protons were identified from 2D ^{15}N -HSQC spectra recorded 12 hours after the $^1\text{H}_2\text{O}$ buffer was exchanged into a $^2\text{H}_2\text{O}$. Detailed descriptions of these experiments along with their original references have been reviewed elsewhere (Cavanagh, 1996; Clore and Gronenborn, 1994). All NMR data were processed using the NMRpipe software system (Delaglio *et al.*, 1995) and analyzed with the program NMRview (Johnson, 1994).

Structure calculations

Structures were calculated with a distance geometry/simulated annealing protocol in the X-PLOR v 3.851 (Brünger, 1996). Initial structure calculations were performed without any assumptions about zinc coordination. Initial structures indicated a cross-brace zinc ligation scheme (see Results and Discussion). Subsequent structure calculations included zinc atoms with additional distance and angle constraints to maintain the tetrahedral bonding geometry of the sites and appropriate bond lengths as previously described by (Neuhaus *et al.*, 1992). Other alternative ligation schemes were tested but were unable to satisfy the experimentally derived

constraints. Structures from these calculations were used for an automated, iterative assignment of remaining NOEs using ARIA (Nilges, 1995; Nilges *et al.*, 1997). In addition to NOE data, calculations used a total of 6 hydrogen bonds and 20 ϕ -dihedral restraints. Quality of the final structures were assessed with PROCHECK-NMR (Laskowski *et al.*, 1996). Figures were prepared using the program PREPI.

Sequence and structure comparisons

Structural comparisons and RMS difference calculations between the PHD, PML (1bor) and IEEHV (1chc) were performed using the least squares fitting algorithm LSQFIT. The graphical overlays were viewed using PREPI. PREPI and LSQFIT were kindly provided by S. Islam and M. Sternberg, Imperial Cancer Research Fund. Sequence alignments were done using Clustal W 1.7.

Acknowledgments. We are grateful for critical reading of the manuscript by Aneel Aggarwal, Graeme Carlile, Alex Kentsis, Pauline McIntosh and Larry Shapiro. We thank Lei Zeng for assistance in setting up NMR experiments. PREPI and LSQFIT programs were a kind gift of Suhail Islam and Mike Sternberg (Imperial Cancer Research Fund). A.D.C. is a NCI trainee. K.L.B.B. is a scholar of the Leukemia Society. Financial support was provided by NIH RO1 CA 80728-01.

Figure Legends

Figure 1. (A) The consensus sequences which define the PHD, RING, and LIM domains. C indicates cysteine, H histidine, and X indicates any residue. Above the consensus sequence is given the number of each metal-ligand, *ml*. In dark blue are the first and second pair of

sequential metal ligands, while in magenta are the third and fourth pair. (B) Demonstration of zinc ligation patterns found in RING and LIM domains. The RING uses a cross-brace ligation scheme while the LIM uses a sequential ligation scheme. Numbers correlate to the number of the metal ligand as defined in A. Zinc atoms are represented by grey ovals and the zinc-binding sites are denoted by roman numerals. (C) Below are ribbon diagrams of the three-dimensional structures. Zinc atoms are represented by spheres. Structures are from the RING of PML (1BOR) and for LIM (1A7I).

Figure 2. (A) Circular dichroism spectrum of untagged KAP-1 PHD prior to addition of EDTA (solid) and after addition (dashed). (B and C) $^1\text{H}/^{15}\text{N}$ -HSQC of KAP-1 PHD produced in minimal media not supplemented with zinc (B) or supplemented with zinc (C).

Figure 3. The KAP-1 PHD structure. (A) α -Carbon overlay of ten KAP-1 PHD domain structures. The residues N-terminal to I627 and C terminal to H670 are disordered and not shown. The average position of the two zinc atoms are represented by white spheres. The blue portion of the ensemble represents the flexible hinge region described in the text. Structural statistics are given in Table I. (B) All atom view of KAP-1 PHD in the same orientation as in A. Main-chain is colored in gray. Side-chains are colored as follows: hydrophobic in yellow, polar (non-charged) in cyan, polar (acidic) in red, and polar (basic) in blue. (C and D) Ribbon diagram of the KAP-1 PHD domain. The β -strands are shown as arrows and zinc ligands in green. (E) Sequence alignment of PHD domains and a schematic depicting zinc ligation.

Figure 4. Comparison of the KAP-1 PHD, PML RING and IEEHV RING. (A) Superposition of KAP-1 PHD (with blue β -strands, a.a. 627-652) and PML RING (in magenta β -strands, a.a. 56 to 76) from the first metal-ligand to the sixth metal-ligand. The white spheres represent zinc atoms with the upper zinc atom being site I. (B) Superposition of zinc-binding site I for KAP-1

(627-632, 647-652), PML (56-71, 76-81), and IEEHV (7-12, 28-33). The metal ligands are colored according to protein; KAP-1 in blue, PML in magenta and IEEHV in green. (C) Superposition of the conserved hydrophobic core residues N-terminal to metal-ligand 5 and C-terminal to metal-ligand 6 (L76 and L81 in PML, F28 and I33 in IEEHV, F647 and H652 in PHD). The side-chains are colored as in B. The core residues from PHD are noted for clarity. The conserved tryptophan within the PHD family (W664 in KAP-1) is seen here inserting between the other core residues, repositioning the core.

Figure 5. Amino acid sequence alignment of the PHD, RING, and LIM families. Site I metal ligands are colored in magenta and site II metal ligands in blue. Conserved hydrophobic core residues are colored in yellow. Accession and PDB codes: PML (1bor), IEEHV (1chc), RAG1 (1rmd), Crp1LIM1 & LIM2 (1B8T), CRIPrat (1IML), KAP-1 (x99644), TIF1 β (s78219), Mi-2 α (x86691), HuATRX (u72937), BARD (NP_00456), DNMTL (AF194032), BRCA1 (A58881), ING1 (AAF37421).

Figure 6. Mutations in the PHD domain of KAP-1 significantly impair the intrinsic repression activity of KAP-1. (A) Schematic diagram illustrating the reporter plasmid (5X-GAL4-UAS-TK-Luciferase) and effector plasmid (GAL4-KAP-1₆₁₉₋₈₃₅). Primary amino acid sequence of the minimal KAP-1 finger is illustrated. The pound (#) symbol represents conserved amino acids that were mutated. Dollar signs (\$) represent non-conserved amino acids in KAP-1 which were mutated to match the corresponding amino acids in TIF1 α /TIF1 γ . Each mutation was made in the context of a GAL4-KAP-1 (619-835) expression construct. (B) Mutations in the KAP-1 PHD domain disrupt its role in transcriptional repression. All experiments were done in NIH/3T3 cells with 5 μ g of the indicated GAL4-fusion protein and 1 μ g of a 5 \times -GAL4(UAS)-TK-luciferase reporter. A black bar represents wild-type, grey bars represent mutations of conserved

amino acid residues, and white bars represent substitutions at non-conserved amino acids. (C) Stable expression of each protein was determined via transfection into COS1 cells followed by immunoprecipitation of ³⁵S-met labeled whole cell extracts with anti-GAL4 (DBD) antiserum (1 µg).

References

Aapola, U., Kawasaki, K., Scott, H.S., Ollila, J., Vihinen, M., Heino, M., Shintani, A., Minoshima, S., Krohn, K., Antonarakis, S.E., Shimizu, N., Kudoh, J. and Peterson, P. (2000) Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family *Genomics*, **65**, 293-298.

Aasland, R., Gibson, T.J. and Stewart, A.F. (1995) The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem Sci*, **20**, 56-59.

Barlow, P.N., Luisi, B., Milner, A., Elliott, M. and Everett, R. (1994) Structure of the C3HC4 domain by 1H-nuclear magnetic resonance spectroscopy. A new structural class of zinc-finger. *J Mol Biol*, **237**, 201-211.

Bock, R. (1979) *Decomposition Methods in Analytical Chemistry*. T & A Constable Ltd, Edinburgh.

Borden, K.L.B. (1998) RING fingers and B-boxes: zinc-binding protein-protein interaction domains. *Biochem Cell Biol*, **76**, 351-358.

Borden, K.L.B. (2000) RING domains: Master Builders of Macromolecular Scaffolds? *J. Mol. Biol.* **295** 1103-1112.

Borden, K.L.B., Boddy, M.N., Lally, J., O'Reilly, N.J., Martin, S., Howe, K., Solomon, E. and Freemont, P.S. (1995a) The solution structure of the RING finger domain from the acute promyelocytic leukaemia proto-oncoprotein PML. *EMBO J*, **14**, 1532-1541.

Borden, K.L.B. and Freemont, P.S. (1996) The RING finger domain: a recent example of a sequence-structure family. *Curr Opin Struct Biol*, **6**, 395-401.

Borden, K.L.B., Lally, J.M., Martin, S.R., O'Reilly, N.J., Etkin, L.D. and Freemont, P.S. (1995b) Novel topology of a zinc-binding domain from a protein involved in regulating early Xenopus development. *EMBO J*, **14**, 5947-5956.

Brehm, A., Nielsen, S.J., Miska, E.A., McCance, D.J., Reid, J.L., Bannister, A.J. and Kouzarides, T. (1999) The E7 oncoprotein associates with Mi2 and histone deacetylase activity to promote cell growth. *EMBO J*, **18**, 2449-2458.

Brunger, A. (1996) *XPLOR (Version 3.843): A system for X-ray Crystallography and NMR*. Yale University Press. Yale University Press, Yale University Press.

Cavanagh, J., Fairbrother, W.J., Palmer, A.G.I., Skelton, N.J. (1996) *Protein NMR Spectroscopy*. Academic Press, San Diego, CA.

Clore, G.M. and Gronenborn, A.M. (1994) Multidimensional heteronuclear nuclear magnetic resonance of proteins. *Methods Enzymol*, **239**, 349-363.

The Finnish-German APECED Consortium. (1997) An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. The Finnish-German APECED Consortium. Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy. *Nat Genet*, **17**, 399-403.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*, **6**, 277-293.

Flores, T.P., Orengo, C.A., Moss, D.S. and Thornton, J.M. (1993) Comparison of conformational characteristics in structurally similar protein pairs. *Protein Sci*, **2**, 1811-1826.

Gibbons, R.J., Bachoo, S., Picketts, D.J., Aftimos, S., Asenbauer, B., Bergoffen, J., Berry, S.A., Dahl, N., Fryer, A., Keppler, K., Kurosawa, K., Levin, M.L., Masuno, M., Neri, G., Pierpont, M.E., Slaney, S.F. and Higgs, D.R. (1997) Mutations in transcriptional regulator ATRX establish the functional significance of a PHD-like domain. *Nat Genet*, **17**, 146-148.

Dawid, I.B., R Toyama. (1994) LIM domains: multiple roles as adapters and functional modifiers in protein interactions. *Trends in Genetics*, **10**. 156-162.

Gunduz, M., Ouchida, M., Fukushima, K., Hanafusa, H., Etani, T., Nishioka, S., Nishizaki, K. and Shimizu, K. (2000) Genomic structure of the human ING1 gene and tumor-specific mutations detected in head and neck squamous cell carcinomas. *Cancer Res*, **60**, 3143-3146.

Jacobson, S. and Pillus, L. (1999) Modifying chromatin and concepts of cancer. *Curr Opin Genet Dev*, **9**, 175-184.

Johnson, M., Correia, JJ Yphantis, DA, Halvorson, HR. (1994) NMR VIEW- a computer program for the visualization and analysis of NMR data. *J Biomol. NMR*, **6**, 603-614.

Kentsis, A. and Borden, K. (2000) Construction of Macromolecular Assemblages in Eukaryotic Processes and their Role in Human Disease: Linking RINGs together. *Current Peptide and Protein Science*, **1**, 49-74.

Klevit, R.E. (1991) Recognition of DNA by Cys2, His2 zinc fingers. *Science*, **253**, 1367-1393.

Laskowski, R.A., Rullmann, J.A., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR*, **8**, 477-486.

Linder, B., Newman, R., Jones, L.K., Debernardi, S., Young, B.D., Freemont, P., Verrijzer, C.P. and Saha, V. (2000) Biochemical analyses of the AF10 protein: the extended LAP/PHD-finger mediates oligomerisation. *J Mol Biol*, **299**, 369-378.

Lu, X., Meng, X., Morris, C.A. and Keating, M.T. (1998) A Novel Human Gene, WSTF, is Deleted in Williams Syndrome. *Genomics*, **54**, 241-249.

Neuhaus, D., Nakaseko, Y., Schwabe, J.W. and Klug, A. (1992) Solution structures of two zinc-finger domains from SWI5 obtained using two-dimensional ^1H nuclear magnetic resonance spectroscopy. A zinc-finger structure with a third strand of beta-sheet. *J Mol Biol*, **228**, 637-651.

Nicholls, A., Sharp, K.A. and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281-296.

Nilges, M. (1995) Calculation of protein structures with ambiguous distance restraints. Automated assignment of ambiguous NOE crosspeaks and disulphide connectivities. *J Mol Biol*, **245**, 645-660.

Nilges, M., Macias, M.J., O'Donoghue, S.I. and Oschkinat, H. (1997) Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from beta-spectrin. *J Mol Biol*, **269**, 408-422.

Omichinski, J.G., Clore, G.M., Schaad, O., Felsenfeld, G., Trainor, C., Appellà, E., S.J., S. and Gronenborn, A.M. (1993) NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science*, **261**, 438-446.

Rinderle, C., Christensen, H.M., Schweiger, S., Lehrach, H. and Yaspo, M.L. (1999) AIRE encodes a nuclear protein co-localizing with cytoskeletal filaments: altered sub-cellular distribution of mutants lacking the PHD zinc fingers. *Hum Mol Genet*, **8**, 277-290.

Ryan, R.F., Schultz, D.C., Ayyanathan, K., Singh, P.B., Friedman, J.R., Fredericks, W.J. and Rauscher, F.J., 3rd. (1999) KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol Cell Biol*, **19**, 4366-4378.

Wallace. (1981) *Analytical Methods for Inductively Coupled Spectrometry*. Perkin Elmer Corp., Norwalk.

Zheng, N., Wang, P, Jeffrey, PD, Pavletich, NP. (2000) Structure of a c-Cbl-UbcH7 Complex: RING domain function in ubiquitin protein ligases. *Cell*, **102**, 533-539.

Table I. Structural statistics

Restraints for structure calculation (residues 627-670)	
Total Restraints	385
Total NOE restraints	317
Intra-residue	78
Sequential	118
Medium range ($2 \leq i-j \leq 4$)	26
Long-range ($ i-j \geq 5$)	95
Dihedral angle ϕ	20
H-bond restraints	12
Zinc distance restraints	20
Zinc angle restraints	16
Final energies (kcal/mol)	
E_{total}	37.1 ± 1.1
E_{NOE}	0.30 ± 0.21
E_{cdih}	0.05 ± 0.04
Coordinate precision (627-653, 664-670)	
R.m.s.d. of backbone atoms	$0.73 \pm 0.16 \text{ \AA}$
R.m.s.d. of all heavy atoms	$1.48 \pm 0.18 \text{ \AA}$
Procheck analysis (627-670)	
Most favoured regions	36.0%
Additional allowed	45.9%
Generously allowed	14.9%
Disallowed region	2.4%

Table II. PHD mutations and structural consequences

Mutations	Location in KAP-1 PHD structure	Transcriptional repression ^a (%)	Structural consequences
CC628,631AA	Zn ²⁺ -ligands of site I	17	loss of metal binding, incorrect folding
G635A	Surface, extended region before β ₁	87	no obvious structural effect
D636A	Surface, extended region before β ₁	74	change in surface charge
M639C	β ₁ , adjacent to metal-ligand of site II (C640)	43	extra zinc ligand, may disrupt site II formation
C643A	Zn ²⁺ -ligands of site II	13	loss of metal binding, incorrect folding
C646V	β ₂ , surface exposed	56	not a zinc ligand, no obvious structural effect
F647Y	Core residue, β ₂	100	conservative substitution
H648C	Zn ²⁺ -ligand of site I	70	Conservative substitution
H648A	Zn ²⁺ -ligand of site I	13	loss of metal binding, incorrect folding
P660A	Flexible hinge	86	no obvious structural effect
W664A	Core residue, β ₃	22	disrupts hydrophobic core
C666A	Zn ²⁺ -ligand of site II	12	loss of metal binding, incorrect folding
SL 667,668 TF	Surface, site II	68	no obvious structural effect
ATRX ^{b,c}			
C628R	Zn ²⁺ -ligand of site I	15	loss of metal binding, incorrect folding
V630S	Surface, site I	85	maybe important for protein-protein interactions
C651F	Zn ²⁺ -ligand of site I	12	loss of metal binding, incorrect folding
P654C	Flexible hinge	42	extra zinc ligand causing incorrect folding
Q657E	Surface, flexible hinge	63	change in surface charge
ING1 ^{b,d}			
C631S	Zn ²⁺ -ligand of site I	---	loss of metal binding, incorrect folding
N632S	Surface, adjacent to metal ligand	---	extra zinc ligand, may disrupt site I formation

^aPercentage compared to transcriptional repression by wild-type KAP-1 PHD

^bnumbers refer to the residue numbering for KAP-1

^cresults for ATRX mutations appear elsewhere (Schultz et al., submitted)

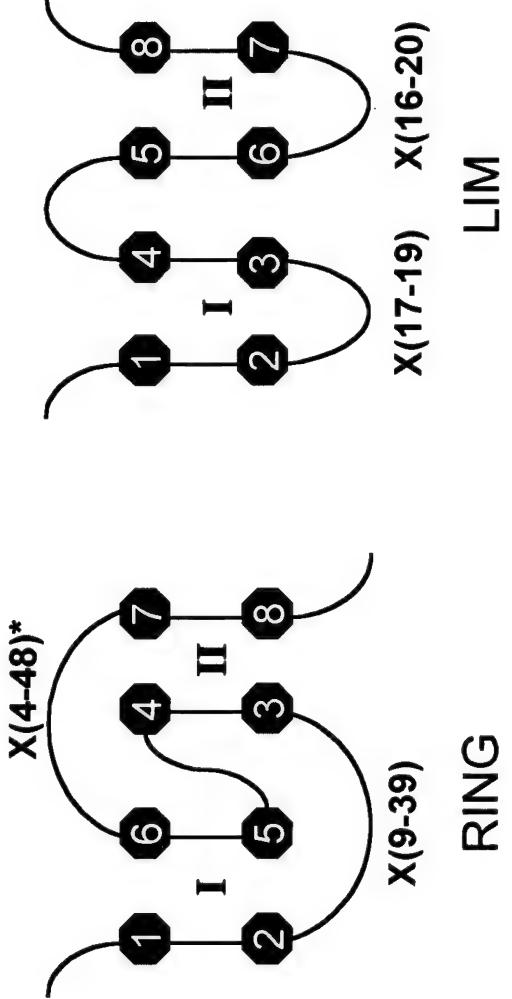
^dtranscriptional assays were not conducted for ING1 mutations

		Metal ligands							
		1	2	3	4	5	6	7	8
C-X ₂ -C-X(09-21)-	C-X(2-4)-	C-X(4-5)-	H-X ₂ -	C-X(12-46)-	C-X ₂ -C				
C-X ₂ -C-X(09-39)-	C-X(1-3)-	H-X(2-3)-	C-X ₂ -	C-X(04-48)-	C-X ₂ -C				
C-X ₂ -C-X(17-19)-	H-X ₂ -	-C-X ₂ -	C-X ₂ -	C-X(16-20)-	C-X ₂ -C				

A

PHD
RING
LIM

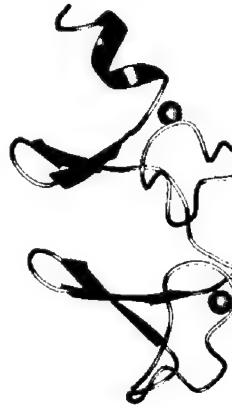
2



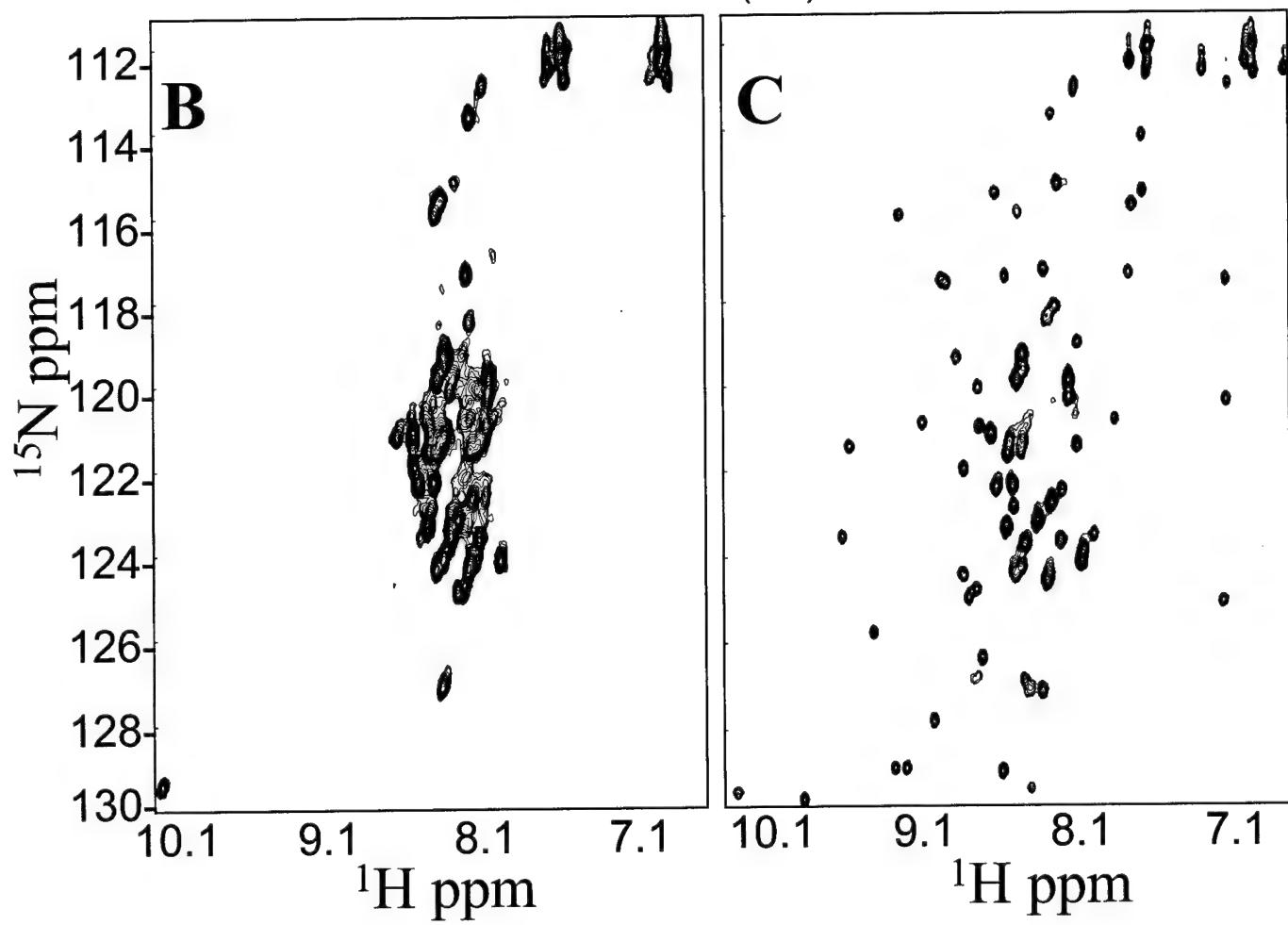
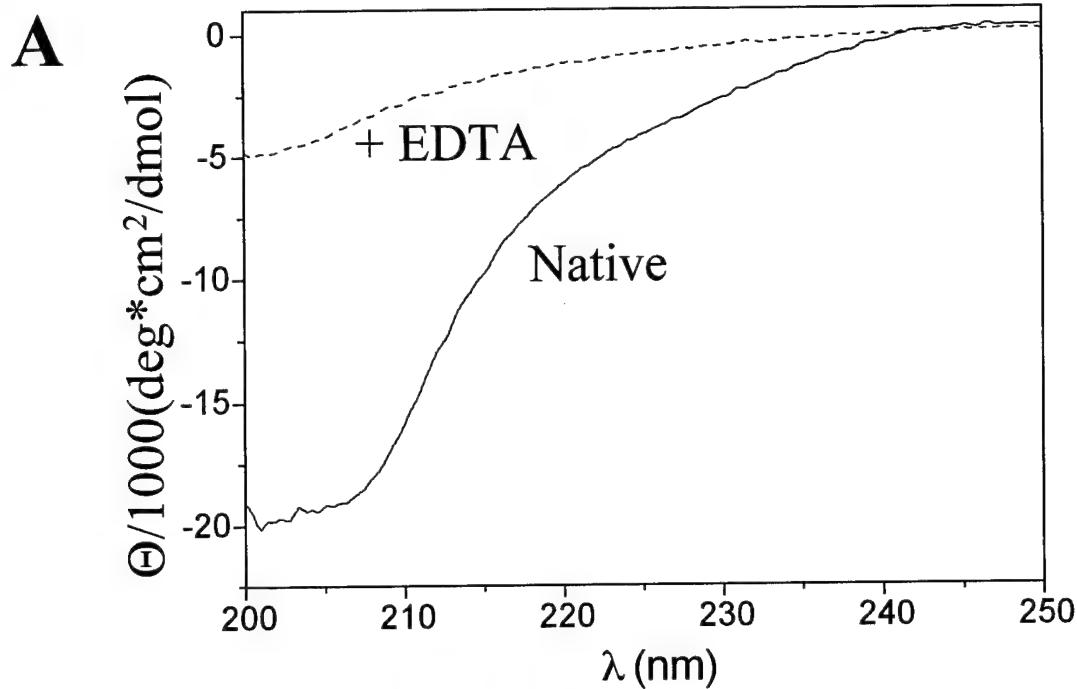
LIM

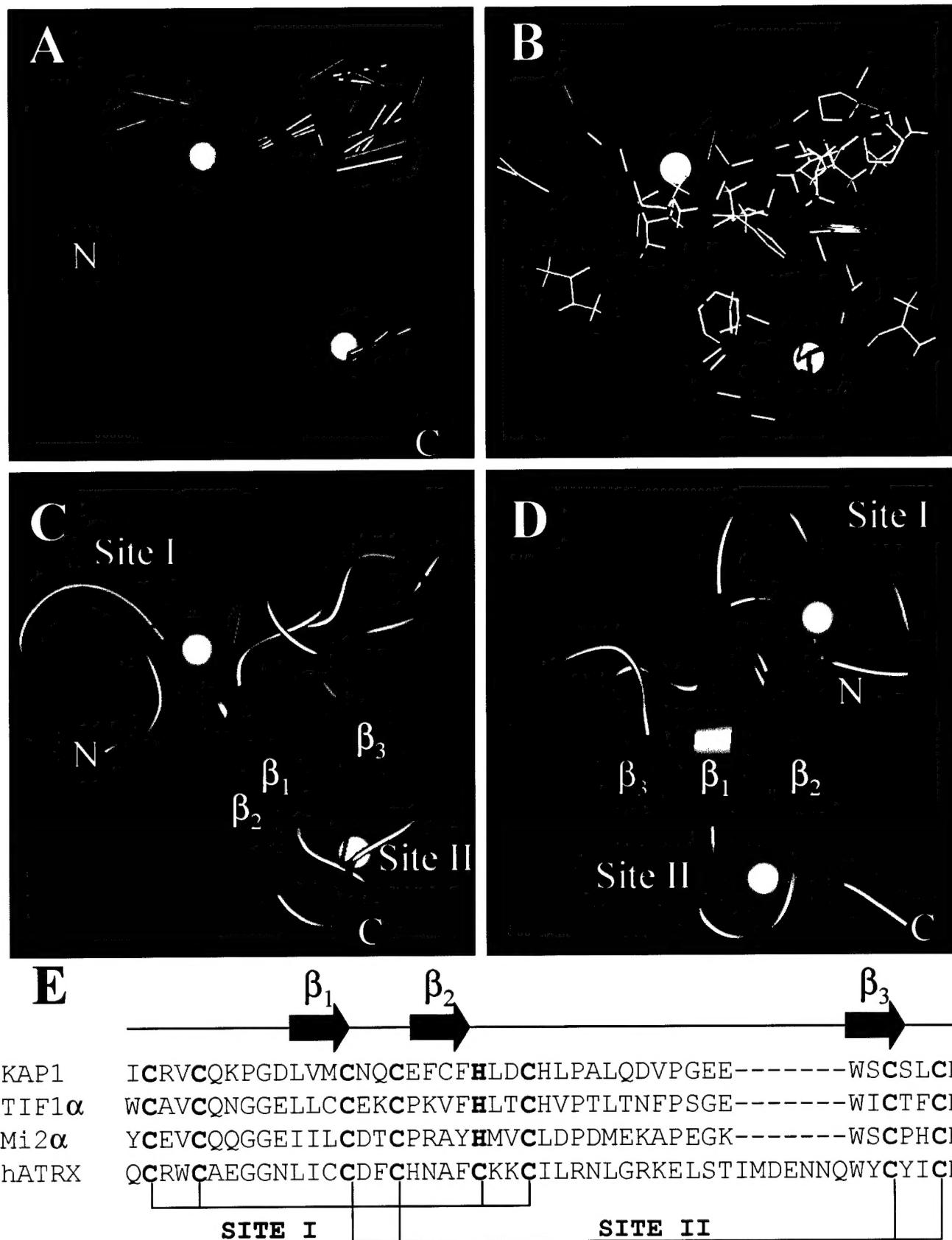
X(17-19) X(16-20)

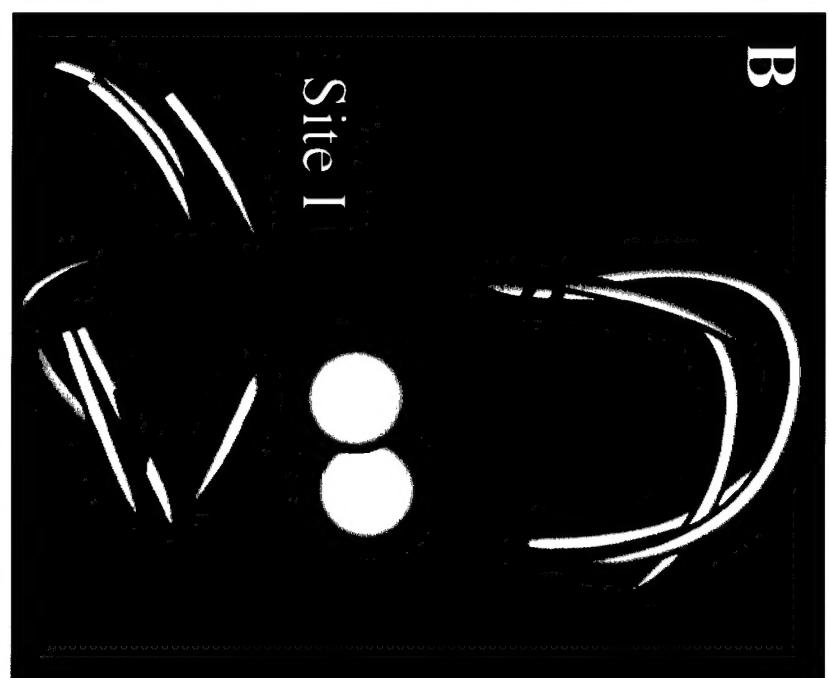
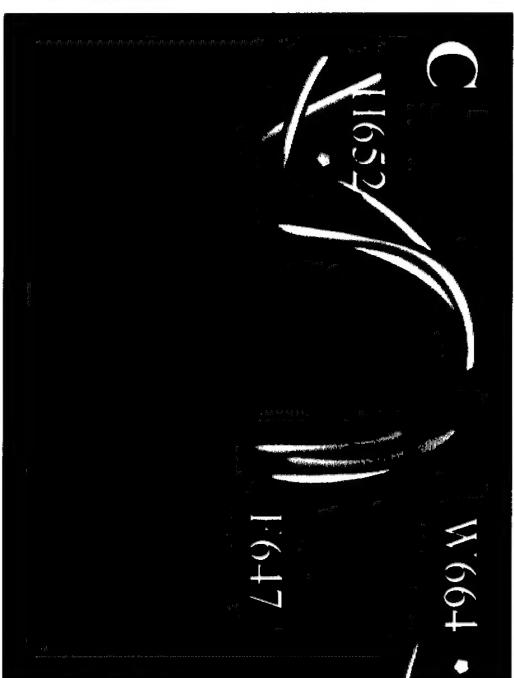
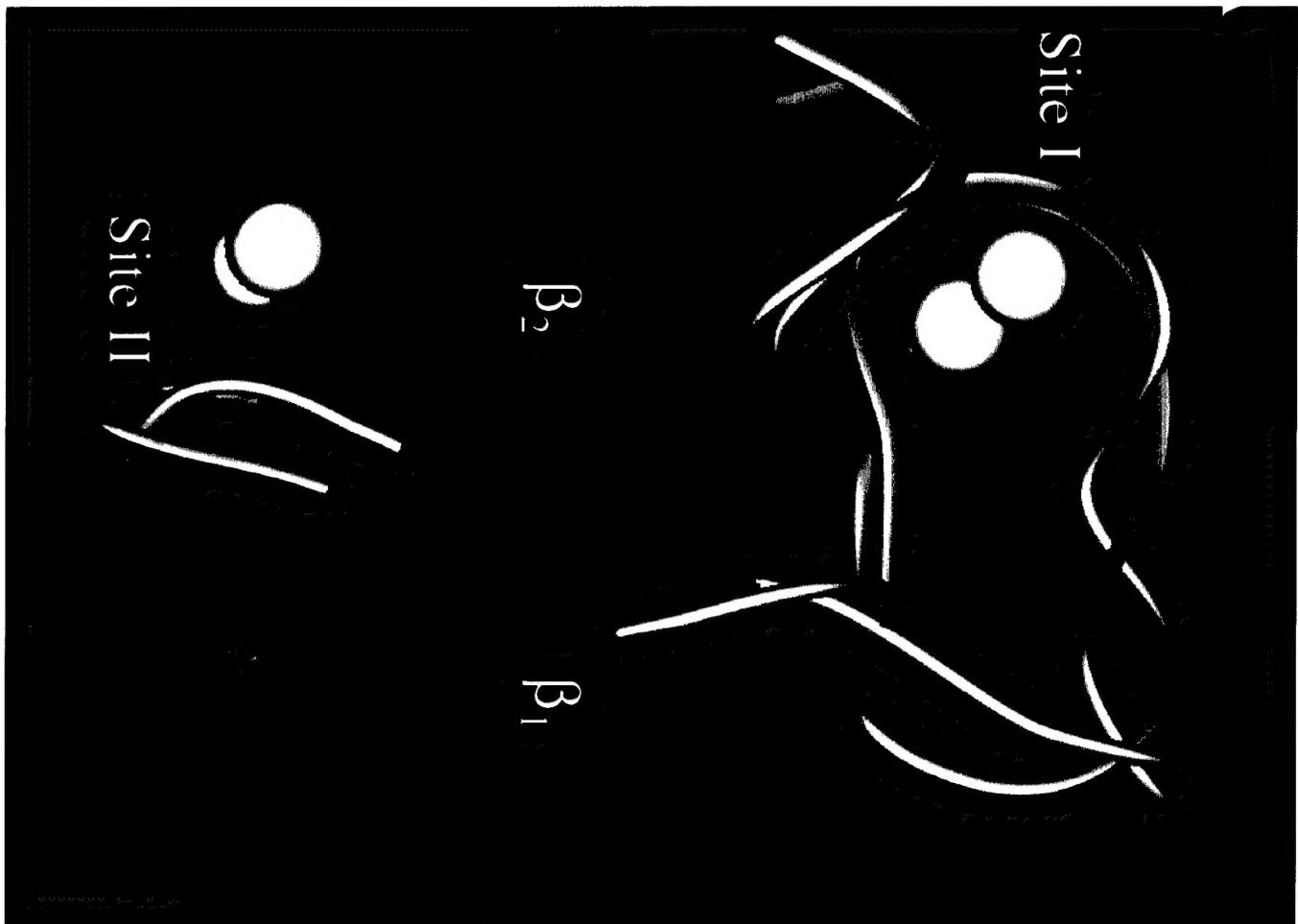
RING



U







PHD

KAP-1	(627) ICRVCQKPGDLVM-----	CNQ--C-EFCFHLD CH ELPALQDVPGEE-----	W	SCS	LCH
TIF1 α	(795) WCAVCQNGGELL-----	CEK--CP-KVFHLS CH EVP <small>T</small> LTNFPSGE-----	W	I	CT FCR
Mi 2 α	(371) YCEVCQQGGEIIL-----	CDT--CP-RAYHM V CILD D PDMEKAPEGK-----	W	S	CP HCE
hATRX	(103) QCRWCAEGGNLIC-----	CDF--C-HNAF V KKC C ILRNL G RKELSTIMDENNNQWYCYICH			
DNMT3A	(93) YCSICCSGETLLI-----	CGNPDC--TRCYC F E C VDSL V GPGTSGKVHAMSN-WV C YLCL			
ING1	(212) YC-LCNQVSYGEMIG-----	CDNDECPIEW F HFS C VGLNHNPKKGK-----	W	Y	CP KCR

RING

PML	(56) R C QQCQA E AKCPKLLP-----	CL-----H-----TL C SG C LEASG-----		M	QC	PICQ
IEEHV	(7) RCPICLEDPSNSYSMALP-----	CL-----H-----AF C YY V C I TRWIRQN-----		P	T	CPLICK
RAG1	(292) SCQICEH I LLADPVETN-----	CK-----H-----VF C RIC I LLRCL V VMG-----		S	Y	CPSCR
BRCA1	(23) EC P ICLE L IKEPVSTK-----	CD-----H-----IF C K F C M L K LLNQKKGP-----		S	O	CP LICK
BARD1	(49) RCSRCTN I LLREPVC L GG-----	CE-----H-----IF C SNC V SDCIG-----		T	G	CP VCY

LIM

Crp1Lim1 (9)	KCG V CQKAVYFAEEVQCEGSSE H KS-----	C FL---C M V C KK N L DSTTVAVHGDE-----	I	Y	C KS-CY
CRPLIM2 (117)	GC P RC G QAVYAAEKV I GAGK S W H KS-----	C FR---CA K C G K S LEST T LADKDGE-----	I	Y	C KG-CY
CRIPrat (2)	KCP K CD K EVYFAERVTSLGKD W HRP-----	C LK---CE K C G K T LTSGGHAEHEGK-----	P	Y	C NHPCY

